

Seq-18 2 w/ Inters

OM protein - protein search, using sw model
Run on: December 17, 2004, 21:53:48 ; Search time 466 Seconds
(without alignments)
1581.157 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITPIYPSGKLH.....DGKLLITVDPAVNGSVIG 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata1/paa/PCTUS COMB.pcp.*
- 2: /cgn2_6/ptodata1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata1/paa/US081 COMB.pcp.*
- 6: /cgn2_6/ptodata1/paa/US082 COMB.pcp.*
- 7: /cgn2_6/ptodata1/paa/US083 COMB.pcp.*
- 8: /cgn2_6/ptodata1/paa/US084 COMB.pcp.*
- 9: /cgn2_6/ptodata1/paa/US085 COMB.pcp.*
- 10: /cgn2_6/ptodata1/paa/US086 COMB.pcp.*
- 11: /cgn2_6/ptodata1/paa/US087 COMB.pcp.*
- 12: /cgn2_6/ptodata1/paa/US088 COMB.pcp.*
- 13: /cgn2_6/ptodata1/paa/US089 COMB.pcp.*
- 14: /cgn2_6/ptodata1/paa/US090 COMB.pcp.*
- 15: /cgn2_6/ptodata1/paa/US091 COMB.pcp.*
- 16: /cgn2_6/ptodata1/paa/US092 COMB.pcp.*
- 17: /cgn2_6/ptodata1/paa/US093 COMB.pcp.*
- 18: /cgn2_6/ptodata1/paa/US094 COMB.pcp.*
- 19: /cgn2_6/ptodata1/paa/US095 COMB.pcp.*
- 20: /cgn2_6/ptodata1/paa/US096 COMB.pcp.*
- 21: /cgn2_6/ptodata1/paa/US097A COMB.pcp.*
- 22: /cgn2_6/ptodata1/paa/US097B COMB.pcp.*
- 23: /cgn2_6/ptodata1/paa/US098 COMB.pcp.*
- 24: /cgn2_6/ptodata1/paa/US099A COMB.pcp.*
- 25: /cgn2_6/ptodata1/paa/US099B COMB.pcp.*
- 26: /cgn2_6/ptodata1/paa/US100 COMB.pcp.*
- 27: /cgn2_6/ptodata1/paa/US101 COMB.pcp.*
- 28: /cgn2_6/ptodata1/paa/US102 COMB.pcp.*
- 29: /cgn2_6/ptodata1/paa/US103 COMB.pcp.*
- 30: /cgn2_6/ptodata1/paa/US104 COMB.pcp.*
- 31: /cgn2_6/ptodata1/paa/US105 COMB.pcp.*
- 32: /cgn2_6/ptodata1/paa/US106 COMB.pcp.*
- 33: /cgn2_6/ptodata1/paa/US107 COMB.pcp.*
- 34: /cgn2_6/ptodata1/paa/US108 COMB.pcp.*
- 35: /cgn2_6/ptodata1/paa/US109 COMB.pcp.*
- 36: /cgn2_6/ptodata1/paa/US160 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3479	100.0	665	1	PCT-US97-06555-2
2	3479	100.0	665	1	Sequence 2, Appli
3	3479	100.0	665	18	Sequence 2, Appli
4	3479	100.0	665	32	Sequence 2, Appli
5	3446	99.1	665	30	Sequence 1498, Ap
6	3446	99.1	679	1	Sequence 13418, A
7	3446	99.1	679	23	Sequence 13416, A
10	3429	98.6	679	28	Sequence 13418, A
11	2883.5	82.9	666	28	Sequence 73893, A
12	2883.5	82.9	666	30	Sequence 74415, A
13	2883	82.9	669	28	Sequence 2228, Ap
14	2732.5	78.5	665	30	Sequence 72423, A
15	2732.5	78.5	682	30	Sequence 10044, A
16	2295	66.0	669	28	Sequence 2226, Ap
17	2291.5	65.9	674	30	Sequence 57809, A
18	2291.5	65.9	674	30	Sequence 6201, Ap
19	2288	65.8	677	1	Sequence 6201, Ap
20	2288	65.8	677	23	Sequence 10663, A
21	2288	65.8	677	26	Sequence 10663, A
22	2288	65.8	677	28	Sequence 42475, A
23	2284.5	65.7	666	1	Sequence 4933, Ap
24	2284.5	65.7	666	23	Sequence 4933, Ap
25	2284.5	65.7	666	26	Sequence 4933, Ap
26	2284.5	65.7	666	35	Sequence 4701, Ap
27	2154	61.9	680	1	Sequence 2080, Ap
28	2154	61.9	680	31	Sequence 2080, Ap
29	2113	60.7	664	28	Sequence 60385, A
30	2040	58.6	652	1	Sequence 2078, Ap
31	2040	58.6	652	31	Sequence 2078, Ap
32	1963.5	56.4	664	22	Sequence 97285, A
33	1949	56.0	659	28	Sequence 71524, A
34	1935.5	55.6	669	28	Sequence 46123, A
35	1934	55.6	657	1	Sequence 12135, A
36	1934	55.6	657	23	Sequence 12135, A
37	1934	55.6	657	26	Sequence 12135, A
38	1931	55.5	657	11	Sequence 12135, A
39	1931	55.5	657	13	Sequence 2, Appli
40	1931	55.5	657	17	Sequence 2, Appli
41	1931	55.5	657	28	Sequence 4393, A
42	1931	55.5	657	34	Sequence 559, App
43	1931	55.5	657	36	Sequence 559, App
44	1930	55.5	657	1	Sequence 13013, A
45	1930	55.5	657	23	Sequence 13013, A

ALIGNMENTS

RESULT 1
PCT-US97-06555-2
; Sequence 2, Application PC/TUS9706555
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/06555
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607999.1
FILING DATE: 16-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US97-06555-2

Query Match 100.0%; Score 3479; DB 1; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDDEHGQIKQK 60
DB 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDDEHGQIKQK 60

QY 61 ABEAGITPOAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAODDIYL 120
DB 61 ABEAGITPOAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAODDIYL 120

QY 121 GEYSGWYSVDESEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDESEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180

QY 181 LVEFFKAHPEFITPDGRLEMLNFTPEGLDLAVSRFTTWGVPVPSNPKHVVYVWIDA 240
DB 181 LVEFFKAHPEFITPDGRLEMLNFTPEGLDLAVSRFTTWGVPVPSNPKHVVYVWIDA 240

QY 241 LLNATYALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILMLLDVLPRLIAH 300
DB 241 LLNATYALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILMLLDVLPRLIAH 300

QY 301 GWFVMDGKMSKSGNVVPEMLVSEYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYE 360
DB 301 GWFVMDGKMSKSGNVVPEMLVSEYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYE 360

QY 361 LANDLGNLNRITVSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRITVSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSTADPHTHMEAVDYPR 420

QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRQOLASVNSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRQOLASVNSHWQASIRVVAHLIEPFMMETS 480

QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPPLNMEEEELAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPPLNMEEEELAYIKEQMEGNKPAV 540

QY 541 EKWNPDVELKLNKDEIKPDPKVEIRVAEYKEVSKVEGSDKLLQFLDAGDGEDRQI 600
DB 541 EKWNPDVELKLNKDEIKPDPKVEIRVAEYKEVSKVEGSDKLLQFLDAGDGEDRQI 600

QY 601 LSGIAKYNEQELVCKKQVIVANLPRKMKKYSQGMILSAEHDGKLTLLTVDPAVFN 660
DB 601 LSGIAKYNEQELVCKKQVIVANLPRKMKKYSQGMILSAEHDGKLTLLTVDPAVFN 660

QY 661 GSVIG 665
|||||

DB 661 GSVIG 665

RESULT 2
PCT-US97-06555A-2
SEQUENCE 2: Application PC/TUS9706555A
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: Novel Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19046
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq For Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06555A
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607999.1
FILING DATE: 16-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US97-06555A-2

Query Match 100.0%; Score 3479; DB 1; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDDEHGQIKQK 60
DB 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDDEHGQIKQK 60

QY 61 ABEAGITPOAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAODDIYL 120
DB 61 ABEAGITPOAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAODDIYL 120

QY 121 GEYSGWYSVDESEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDESEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180

QY 181 LVEFFKAHPEFITPDGRLEMLNFTPEGLDLAVSRFTTWGVPVPSNPKHVVYVWIDA 240
DB 181 LVEFFKAHPEFITPDGRLEMLNFTPEGLDLAVSRFTTWGVPVPSNPKHVVYVWIDA 240

QY 241 LLNATYALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILMLLDVLPRLIAH 300
DB 241 LLNATYALGAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILMLLDVLPRLIAH 300

QY 301 GWFVMDGKMSKSGNVVPEMLVSEYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYE 360
DB 301 GWFVMDGKMSKSGNVVPEMLVSEYGLDPLRYLMLNLPVSGDGTFTPEDYVGRINYE 360

```
QY 361 LANDGLNLRNRTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDPR 420
DB 361 LANDGLNLRNRTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEETAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEETAYIKEQMEGNKPAV 540
QY 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLRDAGDGEDRQI 600
DB 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLRDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVGGKQIVANLKPRKMMKTVSQMILSAEHDGKLTLLTVDPAVFN 660
DB 601 LSGIAKYYPNEQELVGGKQIVANLKPRKMMKTVSQMILSAEHDGKLTLLTVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 3
US-09-432-695-2
; Sequence 2, Application US/09432695
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/432,695
; FILING DATE: 03-Nov-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,059
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-432-695-2
Query Match 100.0%; Score 3479; DB 18; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSEKNFYITPIYVPSGKLHIGSAYTIIACDVLARYKRLMGYDVFYLTGLDEHGKIQOK 60
DB 1 MSEKNFYITPIYVPSGKLHIGSAYTIIACDVLARYKRLMGYDVFYLTGLDEHGKIQOK 60
QY 61 ABEAGITPOAYVDMAGVGVKELWQLLDISYDKFIRTTDDHEKVAQVFERLLAQDDIYL 120
DB 61 ABEAGITPOAYVDMAGVGVKELWQLLDISYDKFIRTTDDHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSGWTSVDSDEEFTESQLAEVFRDAGNVTGGIAPSGHEVEVWVSESYFLRLSKYQDR 180
DB 121 GEYSGWTSVDSDEEFTESQLAEVFRDAGNVTGGIAPSGHEVEVWVSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNEMLRNFIPEGLDLAVSRITTTGWVPVPSNPKHVVYVWIDA 240
DB 181 LVEFFKAHPEFITPDGRNLNEMLRNFIPEGLDLAVSRITTTGWVPVPSNPKHVVYVWIDA 240
QY 241 LLNTATAGYAQDEHGNFDFKFWNGTVFHMVKGKDLRFSIYWPILLNMLDVKLPDRLIAH 300
DB 241 LLNTATAGYAQDEHGNFDFKFWNGTVFHMVKGKDLRFSIYWPILLNMLDVKLPDRLIAH 300
QY 301 GWFVKDCKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGSDGTFTPEDYVGRINYE 360
DB 301 GWFVKDCKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGSDGTFTPEDYVGRINYE 360
QY 361 LANDGLNLRNRTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDPR 420
DB 361 LANDGLNLRNRTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDPR 420
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
DB 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEETAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEETAYIKEQMEGNKPAV 540
QY 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLRDAGDGEDRQI 600
DB 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVAEVEKSVKVEGSDKLIQFLRDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVGGKQIVANLKPRKMMKTVSQMILSAEHDGKLTLLTVDPAVFN 660
DB 601 LSGIAKYYPNEQELVGGKQIVANLKPRKMMKTVSQMILSAEHDGKLTLLTVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 4
US-10-606-607-2
; Sequence 2, Application US/10606607
; GENERAL INFORMATION:
; APPLICANT: LAWLOR, ELIZABETH JANE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GM10253B-3CID1
; CURRENT APPLICATION NUMBER: US/10/606,607
; PRIOR FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10/187,641
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/025,189
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/432,692
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: 08/844,056
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9607999.1
; PRIOR FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
```

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-606-607-2

Query Match 100.0%; Score 3479; DB 32; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.8e-304;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKPYITPTIYPSGKLHIGSAITTIACDVLYARKLMGVDFYLTGLDEHGQKIQK 60
DB 1 MSEKPYITPTIYPSGKLHIGSAITTIACDVLYARKLMGVDFYLTGLDEHGQKIQK 60
QY 61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
DB 61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSCHYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFRLSKYQDR 180
DB 121 GEYSCHYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNEMLRNTEPGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
DB 181 LVEFFKAHPEFITPDGRNLNEMLRNTEPGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
QY 241 LLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
DB 241 LLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
QY 301 GFWFMKDGKMSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFPTEDYVGRINYE 360
DB 301 GFWFMKDGKMSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFPTEDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDLQASVMSHQAASIRVVAHLIEPPMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDLQASVMSHQAASIRVVAHLIEPPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQRLDAGDGEDROI 600
DB 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQRLDAGDGEDROI 600
QY 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
DB 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 5

US-10-472-928-1498
; Sequence 1498, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026925W0
; CURRENT FILING DATE: 2003-09-26
; CURRENT APPLICATION NUMBER: US/10/472,928
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1498
; LENGTH: 665

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: methionyl-tRNA synthetase (metG)
; OTHER INFORMATION: Cellular location: cytoplasm
; OTHER INFORMATION: Similar to strain R6 sequence 15902740 (0.E+01)
US-10-472-928-1498

Query Match 99.1%; Score 3446; DB 30; Length 665;
Best Local Similarity 99.2%; Pred. No. 1.7e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKPYITPTIYPSGKLHIGSAITTIACDVLYARKLMGVDFYLTGLDEHGQKIQK 60
DB 1 MSEKPYITPTIYPSGKLHIGSAITTIACDVLYARKLMGVDFYLTGLDEHGQKIQK 60
QY 61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
DB 61 ABEAGITPQAYVDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSCHYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFRLSKYQDR 180
DB 121 GEYSCHYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSEESYFRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNEMLRNTEPGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
DB 181 LVEFFKAHPEFITPDGRNLNEMLRNTEPGLEDLAVSRITFTWGVVPVSPNKHVVVWIDA 240
QY 241 LLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
DB 241 LLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVKLPDLRIAH 300
QY 301 GFWFMKDGKMSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFPTEDYVGRINYE 360
DB 301 GFWFMKDGKMSKSGNVVYPEMLVERYGLDPLRYILMRNLPGVSDGTFPTEDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDLQASVMSHQAASIRVVAHLIEPPMMETS 480
DB 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDLQASVMSHQAASIRVVAHLIEPPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQRLDAGDGEDROI 600
DB 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSDKLLQRLDAGDGEDROI 600
QY 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
DB 601 LSGIAXYPNEQELVGKQVQIVANLKPROMKKYVSQGMILSAEHDKGLTLITVDPAVFN 660
QY 661 GSVIG 665
DB 661 GSVIG 665

RESULT 6

PCT-US02-03987-13418
; Sequence 13418, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitza Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028VPC
; CURRENT FILING DATE: 2002-02-02
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/267,636

; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13418

Query Match 99.1%; Score 3446; DB 23; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVLYLTGLDGHGQIKQK 60
DB 15 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVLYLTGLDGHGQIKQK 74
QY 61 AEEAGITPQAYVDGMAGVKGELMQLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 75 AEEAGITPQAYVDGMAGVKGELMQLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 134
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 180
DB 135 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRRTFTWGVVPSPNKHVVYVWIDA 240
DB 195 LVEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRRTFTWGVVPSPNKHVVYVWIDA 254
QY 241 LLNYATGALGAQDEHGNFDKFWNGTVFHMVGKDIIRFHSIYWPILLMMLDVKLPDLRLIAH 300
DB 255 LLNYATGALGAQDEHGNFDKFWNGTVFHMVGKDIIRFHSIYWPILLMMLDVKLPDLRLIAH 314
QY 301 GWFVMDGKMSKSGKNVYVPEMLVERYGLDPLRYLYLNRNLPVGSDDGTFTPDYVGRINYE 360
DB 315 GWFVMDGKMSKSGKNVYVPEMLVERYGLDPLRYLYLNRNLPVGSDDGTFTPDYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 420
DB 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQAASIRVVAHLIEPFWMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQAASIRVVAHLIEPFWMETS 494
QY 481 RAVLTQGLBEVSSLENLSLADFPADVTVAAGTPIFPRLNMBEEIAYIKEQMEGNKPAV 540
DB 495 RAVLTQGLBEVSSLENLSLADFPADVTVAAGTPIFPRLNMBEEIAYIKEQMEGNKPAV 554
QY 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVKESKVEGSKLLQFLDAGDGEDROI 600
DB 555 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVKESKVEGSKLLQFLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKKVQIVANLKPRKMKKYSQGMILSAEHDGKLLTLTVDPAPVN 660

; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; PCT-US02-03987-13418

Query Match 99.1%; Score 3446; DB 1; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVLYLTGLDGHGQIKQK 60
DB 15 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVLYLTGLDGHGQIKQK 74
QY 61 AEEAGITPQAYVDGMAGVKGELMQLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 120
DB 75 AEEAGITPQAYVDGMAGVKGELMQLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYL 134
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 180
DB 135 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRRTFTWGVVPSPNKHVVYVWIDA 240
DB 195 LVEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRRTFTWGVVPSPNKHVVYVWIDA 254
QY 241 LLNYATGALGAQDEHGNFDKFWNGTVFHMVGKDIIRFHSIYWPILLMMLDVKLPDLRLIAH 300
DB 255 LLNYATGALGAQDEHGNFDKFWNGTVFHMVGKDIIRFHSIYWPILLMMLDVKLPDLRLIAH 314
QY 301 GWFVMDGKMSKSGKNVYVPEMLVERYGLDPLRYLYLNRNLPVGSDDGTFTPDYVGRINYE 360
DB 315 GWFVMDGKMSKSGKNVYVPEMLVERYGLDPLRYLYLNRNLPVGSDDGTFTPDYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 420
DB 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQAASIRVVAHLIEPFWMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQAASIRVVAHLIEPFWMETS 494
QY 481 RAVLTQGLBEVSSLENLSLADFPADVTVAAGTPIFPRLNMBEEIAYIKEQMEGNKPAV 540
DB 495 RAVLTQGLBEVSSLENLSLADFPADVTVAAGTPIFPRLNMBEEIAYIKEQMEGNKPAV 554
QY 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVKESKVEGSKLLQFLDAGDGEDROI 600
DB 555 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVKESKVEGSKLLQFLDAGDGEDROI 614
QY 601 LSGIAKYYPNEQELVGKKVQIVANLKPRKMKKYSQGMILSAEHDGKLLTLTVDPAPVN 660
DB 615 LSGIAKYYPNEQELVGKKVQIVANLKPRKMKKYSQGMILSAEHDGKLLTLTVDPAPVN 674
QY 661 GSVIG 665
DB 675 GSVIG 679

RESULT 7
US-09-815-242-13418
; Sequence 13418, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

Db 615 LSGIATYPNEQELVGVKKQIVANLKPROMKKYVSGQMILSAEHDKGLTLLTVDPAVEN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 8
US-10-072-851-13418
; Sequence 13418, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Boone, Charles
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-072-851-13418

Query Match 99.1%; Score 3446; DB 26; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDVFLTGLDHHGKIQOK 60
Db 15 MSEKNYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDVFLTGLDHHGKIQOK 74
QY 61 ABEAGITPQAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 120
Db 75 ABEAGITPQAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIYL 134
QY 121 GEYSGHYSDDEFFTESQLAEVRDEAGNTGGIAPSGHEVWVWSESYFLRLSKYQDR 180
Db 135 GEYSGHYSDDEFFTESQLAEVRDEAGNTGGIAPSGHEVWVWSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPEFTTPDGLNLMERNFTPEGLDLAVSRFTTGWGVPSNPKHVVVWIDA 240
Db 195 LVEFFKAHPEFTTPDGLNLMERNFTPEGLDLAVSRFTTGWGVPSNPKHVVVWIDA 254
QY 241 LLNATATAGYAQDEHGNFDPKFWNGTVFHMVGKDILRFHSIYWPILLMLDVKLPDLIAH 300
Db 255 LLNATATAGYAQDEHGNFDPKFWNGTVFHMVGKDILRFHSIYWPILLMLDVKLPDLIAH 314
QY 301 GNFVMDGKXSKGNVYVPEMLVERYGIDPLRYVLMNLPVSGDTFTPDYVGRINYE 360
Db 315 GNFVMDGKXSKGNVYVPEMLVERYGIDPLRYVLMNLPVSGDTFTPDYVGRINYE 374
QY 361 LANDLGNLNRVTSIMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 420
Db 375 LANDLGNLNRVTSIMINKYFDGQIPAYVEGVTEFDFHVLAEVAEKSIADPHTHMEAVDYPR 434
QY 421 ALEAVNTLISRTNKYIDETAPWVLKDDEALRDLQASVMSHWQASIRVVVAHLIEPFWMTS 480

Db 435 ALEAVNTLISRTNKYIDETAPWVLKDDEALRDLQASVMSHWQASIRVVVAHLIEPFWMTS 494
QY 481 RAVLTQIGLEVSLENLSLADFPADVTVVAKGTPIPPRLNMBEEETAYIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEVSLENLSLADFPADVTVVAKGTPIPPRLNMBEEETAYIKEQMEGNKPAV 554
QY 541 EKWNPDVELKLNKDEIKFEDDKVEIRVAEYKESKVEGSKLLQFRLDAGDGDROI 600
Db 555 EKWNPDVELKLNKDEIKFEDDKVEIRVAEYKESKVEGSKLLQFRLDAGDGDROI 614
QY 601 LSGIATYPNEQELVGVKKQIVANLKPROMKKYVSGQMILSAEHDKGLTLLTVDPAVEN 660
Db 615 LSGIATYPNEQELVGVKKQIVANLKPROMKKYVSGQMILSAEHDKGLTLLTVDPAVEN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 9
US-10-282-122A-73893
; Sequence 73893, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73893
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73893

Query Match 99.1%; Score 3446; DB 28; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.8e-301;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYTLGLDGHGQKIOQ 60
Db 15 MSEKNFYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYTLGLDGHGQKIOQ 74
QY 61 AEEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYL 120
Db 75 AEEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYL 134
QY 121 GEYSWMYSVDSDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
Db 135 GEYSWMYSVDSDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFTTPDGLNEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVVWIDA 240
Db 195 LVEFFKAHPFTTPDGLNEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVVWIDA 254
QY 241 LLNVATAGYAQDEHGNFDFKFWNGTVFHMVGKDLRFHISIYWPILLMMLDVKLPDLRLIAH 300
Db 255 LLNVATAGYAQDEHGNFDFKFWNGTVFHMVGKDLRFHISIYWPILLMMLDVKLPDLRLIAH 314
QY 301 GWFVMDGKMSKSGNVTYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 360
Db 315 GWFVMDGKMSKSGNVTYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLLARTYSMINKYFDGQIPAYVEGVTDFDHVLAEVAEKSIADPHTHMEAVDYPR 420
Db 375 LANDLGNLLARTYSMINKYFDGQIPAYVEGVTDFDHVLAEVAEKSIADPHTHMEAVDYPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHQAASIRVVAHLIEPFMMETS 480
Db 435 ALEAVWTLISRTNKYIDETAPWVLAKDEALRDQLASVMSHQAASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSADPVDVTVVAKGTPIPPRLNMEEBEIAIYKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSADPVDVTVVAKGTPIPPRLNMEEBEIAIYKEQMEGNKPAV 554
QY 541 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGDGEDRQI 600
Db 555 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 10
US-10-640-833-2751
; Sequence 2751, Application US/10640833
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiondong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH03-15
; CURRENT APPLICATION NUMBER: US/10/640,833
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2751

; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-640-833-2751
Query Match 98.6%; Score 3429; DB 32; Length 679;
Best Local Similarity 98.8%; Pred. No. 6.2e-300;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYTLGLDGHGQKIOQ 60
Db 15 MSEKNFYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYTLGLDGHGQKIOQ 74
QY 61 AEEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYL 120
Db 75 AEEAGITPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYL 134
QY 121 GEYSWMYSVDSDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 180
Db 135 GEYSWMYSVDSDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDR 194
QY 181 LVEFFKAHPFTTPDGLNEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVVWIDA 240
Db 195 LVEFFKAHPFTTPDGLNEMLRNFIEPGLDLAVSRITFTWGPVPSNPKHVVVWIDA 254
QY 241 LLNVATAGYAQDEHGNFDFKFWNGTVFHMVGKDLRFHISIYWPILLMMLDVKLPDLRLIAH 300
Db 255 LLNVATAGYAQDEHGNFDFKFWNGTVFHMVGKDLRFHISIYWPILLMMLDVKLPDLRLIAH 314
QY 301 GWFVMDGKMSKSGNVTYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 360
Db 315 GWFVMDGKMSKSGNVTYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLLARTYSMINKYFDGQIPAYVEGVTDFDHVLAEVAEKSIADPHTHMEAVDYPR 420
Db 375 LANDLGNLLARTYSMINKYFDGQIPAYVEGVTDFDHVLAEVAEKSIADPHTHMEAVDYPR 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHQAASIRVVAHLIEPFMMETS 480
Db 435 ALEAVWTLISRTNKYIDETAPWVLAKDEALRDQLASVMSHQAASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSADPVDVTVVAKGTPIPPRLNMEEBEIAIYKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSADPVDVTVVAKGTPIPPRLNMEEBEIAIYKEQMEGNKPAV 554
QY 541 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGDGEDRQI 600
Db 555 EKWNPDVEVLKLNKDEIKFEDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGKVKQIVANLKPRKMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 11
US-10-282-122A-74415
; Sequence 74415, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Jianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUIPA.0344
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74415
; LENGTH: 666
; TYPE: PR1
; ORGANISM: Streptococcus pyogenes
; US-10-282-122A-74415

```

```

Query Match      82.9%; Score 2883.5; DB 28; Length 666;
Best Local Similarity 81.3%; Pred. No. 1.3e-250;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGKIQOK 60
DB 1 MMKKPFVITPIYPYSGKLHIGSAITTIACDVLARYKRLMGHEVFYLTGLDEHGKIQOK 60
QY 61 ABEAGITPQAYDGMAGVKELMQLLDISYDKFIRTTDDYHEKVAAQVFERLLAQQDIYL 120
DB 61 AKEAGITPQTYVDNMAKVKALMQLLDISYDTFIRTTDDYHEVVAAPFEKLLAQQDIYL 120
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDSDEFFTESQLKEVFRDEGQVIGGIAPSGHEVEMVSESYFLRLSKYDDR 180
QY 181 LVFEFFKAHPFIPDGRNEMLNFTPEGLDLAVSRITFTWGPVPSPKHHVYVWIDA 240
DB 181 LVAFFKERPDPIQDGRMNMKNFTPEGLDLAVSRITFTWGPVPSPKHHVYVWIDA 240
QY 241 LLYNATGALGYRQANHANFDKFMNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
DB 241 LLYNATGALGYRQANHANFDKFMNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GFWFMKDGKMSKGNVVPYEMLVVERFGLDPLRYILMRSLPVGSDGTFPTEDYVGRINYE 360
DB 301 GFWFMKDGKMSKGNVVPYEMLVVERFGLDPLRYILMRSLPVGSDGTFPTEDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTBFHVLAEVAKSTADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTBFHVLAEVAKSTADPHTHMEAVDYPR 420
QY 421 ALSAVWTLISRTNKYIDETAPWLDKDEALRQOLASVMSHWQASIRVVAHLIPEPFMETS 480
DB 421 ALSAVWTLISRTNKYIDETAPWLDKDEALRQOLASVMSHWQASIRVVAHLIPEPFMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADPADVTYVAKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADPADVTYVAKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
QY 481 AATMAQUGLEPVSJLSTLALDFFPANTKVAVKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
DB 481 AATMAQUGLEPVSJLSTLALDFFPANTKVAVKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540

```

```

QY 541 --EKWNPDEVELKLNKDEIKFEDFDKVEIRVAEVEKVEKSGDKLQRLDAGDGEDR 598
DB 540 SQEKWNPDEVELKLNKDEIKFEDFDKVEIRVAEVEKVEKSGDKLQRLDAGDGEDR 599
QY 599 QILSGIAKYPNEQELVGVQIVANLKPRKMMKVKYVSOQMILSAEHDGKLTLLTVDPAV 658
DB 600 QILSGIAKYPNEQELVGVQIVANLKPRKMMKVKYVSOQMILSAEHDGKLTLLTVDSSV 659
QY 659 PNGSVIG 665
DB 660 PNGSIIG 666

RESULT 12
US-10-415-182A-2228
; Sequence 2228, Application US/10415182A
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/415,182A
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12024
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 2228
; LENGTH: 666
; TYPE: PR1
; ORGANISM: Streptococcus pyogenes
; US-10-415-182A-2228

```

```

Query Match      82.9%; Score 2883.5; DB 30; Length 666;
Best Local Similarity 81.9%; Pred. No. 1.3e-250;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPIYPYSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGKIQOK 60
DB 1 MMKKPFVITPIYPYSGKLHIGSAITTIACDVLARYKRLMGHEVFYLTGLDEHGKIQOK 60
QY 61 ABEAGITPQAYDGMAGVKELMQLLDISYDKFIRTTDDYHEKVAAQVFERLLAQQDIYL 120
DB 61 AKEAGITPQTYVDNMAKVKALMQLLDISYDTFIRTTDDYHEVVAAPFEKLLAQQDIYL 120
QY 121 GEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSESYFLRLSKYQDR 180
DB 121 GEYSGWYSVDSDEFFTESQLKEVFRDEGQVIGGIAPSGHEVEMVSESYFLRLSKYDDR 180
QY 181 LVFEFFKAHPFIPDGRNEMLNFTPEGLDLAVSRITFTWGPVPSPKHHVYVWIDA 240
DB 181 LVAFFKERPDPIQDGRMNMKNFTPEGLDLAVSRITFTWGPVPSPKHHVYVWIDA 240
QY 241 LLYNATGALGYRQANHANFDKFMNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
DB 241 LLYNATGALGYRQANHANFDKFMNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GFWFMKDGKMSKGNVVPYEMLVVERFGLDPLRYILMRSLPVGSDGTFPTEDYVGRINYE 360
DB 301 GFWFMKDGKMSKGNVVPYEMLVVERFGLDPLRYILMRSLPVGSDGTFPTEDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTBFHVLAEVAKSTADPHTHMEAVDYPR 420
DB 361 LANDLGNLNRVTSMINKYFDGOIPAYVEGVTBFHVLAEVAKSTADPHTHMEAVDYPR 420
QY 421 ALSAVWTLISRTNKYIDETAPWLDKDEALRQOLASVMSHWQASIRVVAHLIPEPFMETS 480
DB 421 ALSAVWTLISRTNKYIDETAPWLDKDEALRQOLASVMSHWQASIRVVAHLIPEPFMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADPADVTYVAKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
DB 481 RAVLTQIGLEEVSSLENLSLADPADVTYVAKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
QY 481 AATMAQUGLEPVSJLSTLALDFFPANTKVAVKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540
DB 481 AATMAQUGLEPVSJLSTLALDFFPANTKVAVKGTPIFFPLNNEEBIAIYKEQVEGNKPAV 540

```

[illegible]

RESULT 13

US-10-282-122A-72423
 ; Sequence 72423, Application US/10282122A
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

```

?  ORGANISM: Streptococcus mutans
US-10-282-122A-73423

Query Match      82.9%; Score 2883; DB 28; Length 669;
Best Local Similarity 82.4%; Pred. N.1.5e-250;
Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;

Qy 1 MSEKN-FYITPTPIYPGSKLHIGAYTTIACDVLARYKRLMGYDVPYLTGLDEHGQIQ 59
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1 MTEQKPYIITPTPIYPGSKLHIGAYTTIACDVLARYKRMHNDVYLTGLDEHGQIQ 60

```

60	QY	KAERAGITPOAVDGMVGVKELWOLLDLSYDKFIRTTDDYHEKVVAQVPERLLAADDIY	119
61	Db	KSEBAGITPOAVDGMVGVKELWOLLDLSYDKFIRTTDDYHEKVVAQVFEKLLAQGDIIY	120
120	QY	LGEYSGWYSVSDDEFTTSQIAEVRDEAGNTGGIAPSGCHEVWVSESYSFURLSKYQD	179
121	Db	LGEYSGWYSVSDDEFTTSQIBEIFRDEBKGVIIGGIAPSGCHEVWVSESYSFURLSNYAD	180
180	QY	RLVEFPFKAPHEFITPDGRNLNMRNFIPEGLEDLAVSRITFTWGVVPVPSNPKHVVVWID	239
181	Db	RLVDFPFAHPDFIQPDGRNNEIINKFIEGLEDLAVSRITFTWGVKVPSPKHVVVWID	240
240	QY	ALLNATATALGYAQDEHGNFDFKFWNGTVPFMVGKDLIRFHSIYWPIILLMMLDVKLPORLIA	299
241	Db	ALLNATATALGYQDAFANFDKFWNGTVPFMVGKDLIRFHSIYWPIILLMMLDLKQPERLIA	300
300	QY	HGWFVMKDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLPVGSDGTFTPEDYVGRINY	359
301	Db	HGWFVMKDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRSLPVGSDGTFTPEDYVGRINY	360
360	QY	ELANDGLNLLNRTVSMINKYFGOGIIPAYVGVGTPEFDHVLAEVAEKSIADPHITHEAVDYP	419
361	Db	ELANDGLNLLNRTVSMINKYFGGEIPAYVENVTDFDADLAADVADNLAKYHKQNAVDYP	420
420	QY	RALEAVTLLISRTNKYIDETAPVWLKDDEALRDQLASVMSHWOASIRVVAHLIEPPFMMET	479
421	Db	RALETVMNIISRTNKYIDETTPWLAKDKAKRDDLAAVMAHLAASLRVVAHLIQPFMMTT	480
480	QY	SRAVLTOIGLEEVSSLENI SLADFADVTVAAGTPIFFRLNMEELIAYIKEQWEGNKPA	539
481	Db	SNAIMQLGLGAQDFDEGLELACLPTGVKVIAGTPIFFRLNDWDEIAYIKEQW-GDNSA	539
540	QY	V----EKWNPPEVELKLNKDEIKFEDFOKPIRVAEVEKSVKVGSDKLLQFRLDAGDG	595
540	Db	IAQEEBKEWNPPTVELKNEKKAIKFDFDOKLEIRVAEVEKSVKVGSDKLLKFLDAGDG	599
596	QY	EDROILSGIAKYPNPBOELVGKVKQIVANLKPRKMMKXYVQSGMILSAEH-DGKLTLLTV	654
600	Db	EDROILSGIAKYPNELDLVGKLLQIVANLKPRKM-GLVSGQMILSAEHADGKLTIVTV	658
655	QY	DPVAPNGSVIG	665
659	Db	DSSVPNGSQIG	669

RESULT 14

```

US-10-415-182A-10044
; Sequence 10044, Application US/10415182A
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/415,182A
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12024
; SOFTWARE: Seqwin99, version 1.02
; SEQ ID NO 10044
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-10-415-182A-10044

```

Query Match 78.5%; Score 2732.5; DB 30; Length 665;
Best Local Similarity 78.2%; Pred. No. 6e-237;
Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;

Query Match 78.5%; Score 2732.5; DB 30; Length 665;

```

QY 3 EKXFIITPIYPSGKHLHIGSAVTTIACDVLARYKELMGVDVFLYTLGLDGHGKIQKAE 62
Db 4 KXSFYITPIYPSGKHLHIGSAVTTIACDVLARYKELMGVDVFLYTLGLDGHGKIQKAE 63
QY 63 EAGITPOAVYDGNMAGVKGELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYIGE 122
Db 64 EAGITPOAVYDGNMAGVKGELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYIGE 123
QY 123 YSGWYSVDEEFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRIV 182
Db 124 YSGWYSVDEEFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRIV 183
QY 183 EPPKAHPEFTPDGRNEMLNFIIEGLEDLAVSRITFTTWGVPVPSNPKHVVYVWIDALL 242
Db 184 AYVAEHPEFTPDGRNEMLNFIIEGLEDLAVSRITFTTWGVPVPSNPKHVVYVWIDALL 243
QY 243 NYATALLGYA-ODEHGNFDFKFWNGTVFHMVKDILRFHSIYWPILLMMLDVKLPDLRIAHG 301
Db 244 NYATALLGYA-ODEHGNFDFKFWNGTVFHMVKDILRFHSIYWPILLMMLDVKLPDLRIAHG 302
QY 302 WFMVQDGKMSKSGNVTYVPEMLVERGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 361
Db 303 WFMVQDGKMSKSGNVTYVPEMLVERGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 362
QY 362 ANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSADFTHMEAVDYPRA 421
Db 363 ANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSADFTHMEAVDYPRA 421
QY 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVVAHLIPEFMMETSR 481
Db 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVVAHLIPEFMMETSR 481
QY 482 AVLTQIGLEEVSSLENLSADFPADVTVVAKGTPIPPRLNMEEEIAYIKEQMEGNKPAYE 541
Db 482 AVLTQIGLEEVSSLENLSADFPADVTVVAKGTPIPPRLNMEEEIAYIKEQMEGNKPAYE 541
QY 542 KEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGDKLQFRLDAGDGEDROL 601
Db 542 KEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGDKLQFRLDAGDGEDROL 601
QY 602 SGIATYPNEQELVGGKQVIVANLKPRKMKKYSQGMILSAHDGKLTLLTVDPAVPG 661
Db 602 SGIATYPNEQELVGGKQVIVANLKPRKMKKYSQGMILSAHDGKLTLLTVDPAVPG 661
QY 662 SVIG 665
Db 662 SVIG 665

```

```

RESULT 15
US-10-415-182A-2226
; Sequence 2226, Application US/10415182A
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/10/415,182A
; PRIOR FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12024
; SOFTWARE: Seqwin99, version 1.02
; SEQ ID NO 2226
; LENGTH: 682
; TYPE: PR
; ORGANISM: Streptococcus agalactiae
US-10-415-182A-2226

```

```

Best Local Similarity 78.2%; Pred. No. 6.2e-237;
Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
QY 3 EKXFIITPIYPSGKHLHIGSAVTTIACDVLARYKELMGVDVFLYTLGLDGHGKIQKAE 62
Db 21 KXSFYITPIYPSGKHLHIGSAVTTIACDVLARYKELMGVDVFLYTLGLDGHGKIQKAE 80
QY 63 EAGITPOAVYDGNMAGVKGELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYIGE 122
Db 81 EAGITPOAVYDGNMAGVKGELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDIYIGE 140
QY 123 YSGWYSVDEEFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRIV 182
Db 141 YSGWYSVDEEFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRIV 200
QY 183 EPPKAHPEFTPDGRNEMLNFIIEGLEDLAVSRITFTTWGVPVPSNPKHVVYVWIDALL 242
Db 201 AYVAEHPEFTPDGRNEMLNFIIEGLEDLAVSRITFTTWGVPVPSNPKHVVYVWIDALL 260
QY 243 NYATALLGYA-ODEHGNFDFKFWNGTVFHMVKDILRFHSIYWPILLMMLDVKLPDLRIAHG 301
Db 261 NYATALLGYA-ODEHGNFDFKFWNGTVFHMVKDILRFHSIYWPILLMMLDVKLPDLRIAHG 319
QY 302 WFMVQDGKMSKSGNVTYVPEMLVERGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 361
Db 320 WFMVQDGKMSKSGNVTYVPEMLVERGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 379
QY 362 ANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSADFTHMEAVDYPRA 421
Db 380 ANDLGNLLNRTVSMINKYFDGQIPAYVEGVTDFHVLAEVAKSADFTHMEAVDYPRA 438
QY 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVVAHLIPEFMMETSR 481
Db 439 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVVAHLIPEFMMETSR 498
QY 482 AVLTQIGLEEVSSLENLSADFPADVTVVAKGTPIPPRLNMEEEIAYIKEQMEGNKPAYE 541
Db 499 AVLTQIGLEEVSSLENLSADFPADVTVVAKGTPIPPRLNMEEEIAYIKEQMEGNKPAYE 558
QY 542 KEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGDKLQFRLDAGDGEDROL 601
Db 559 KEWNPDEVELKLNKDEIKFEDFDKVEIRVAEVKEVSKVSGDKLQFRLDAGDGEDROL 618
QY 602 SGIATYPNEQELVGGKQVIVANLKPRKMKKYSQGMILSAHDGKLTLLTVDPAVPG 661
Db 619 SGIATYPNEQELVGGKQVIVANLKPRKMKKYSQGMILSAHDGKLTLLTVDPAVPG 678
QY 662 SVIG 665
Db 679 SVIG 682

```

Search completed: December 17, 2004, 22:08:59
Job time : 470 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:59:37 ; Search time 27 Seconds
(without alignments)
1110.332 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479

Sequence: 1 MSEKNFYITTPYPSGKLH.....DGKLTLLTVDPVNGSVIG 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 168998 seqs, 45081143 residues

Total number of hits satisfying chosen parameters: 168998

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:

- 1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap.*
- 7: /cgn2_6/ptodata/2/paa/US11_NEW_COMB.pap.*
- 8: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3446	99.1	679	1 PCT-US02-09107B-73893	Sequence 73893, A
2	2883.5	82.9	666	1 PCT-US02-09107B-74415	Sequence 74415, A
3	2883	82.9	669	1 PCT-US02-09107B-72423	Sequence 72423, A
4	2295	66.0	669	1 PCT-US02-09107B-75809	Sequence 75809, A
5	2288	65.8	677	1 PCT-US02-09107B-42475	Sequence 42475, A
6	2113	60.7	664	1 PCT-US02-09107B-60385	Sequence 60385, A
7	2013.5	57.9	658	8 US-60-622-712-354	Sequence 354, App
8	1949	56.0	659	1 PCT-US02-09107B-71524	Sequence 71524, A
9	1935.5	55.6	669	1 PCT-US02-09107B-46123	Sequence 46123, A
10	1931	55.5	657	1 PCT-US02-09107B-43933	Sequence 43933, A
11	1922.5	55.3	656	1 PCT-US02-09107B-71042	Sequence 71042, A
12	1911	54.9	662	1 PCT-US02-37204-4240	Sequence 4240, Ap
13	1678	48.2	674	1 PCT-US02-09107B-52642	Sequence 52642, A
14	1640.5	47.2	644	1 PCT-US02-09107B-51647	Sequence 51647, A
15	1572	45.2	645	1 PCT-US02-09107B-53102	Sequence 53102, A
16	1383	39.8	670	1 PCT-US02-09107B-46354	Sequence 46354, A
17	1180.5	34.2	651	1 PCT-US03-31318-11428	Sequence 11428, A
18	1189.5	34.2	628	1 PCT-US02-09107B-54439	Sequence 54439, A
19	1153.5	33.2	521	1 PCT-US02-09107B-76747	Sequence 76747, A
20	1148.5	33.0	650	1 PCT-US02-09107B-58709	Sequence 58709, A
21	1045	30.0	519	1 PCT-US02-09107B-62338	Sequence 62338, A
22	1042.5	30.0	537	1 PCT-US02-09107B-62166	Sequence 62166, A
23	1040	29.9	537	1 PCT-US02-09107B-63675	Sequence 63675, A
24	1037	29.8	519	1 PCT-US02-09107B-64482	Sequence 64482, A
25	871.5	25.1	512	1 PCT-US02-09107B-64107	Sequence 64107, A

ALIGNMENTS

RESULT 1

PCT-US02-09107B-73893
; Sequence 73893, Application PC/TUS0209107B
; GENERAL INFORMATION:

; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA_034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B

; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993

; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923

; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851

; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699

; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73893

; LENGTH: 679

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

PCT-US02-09107B-73893

Query Match 99.1%; Score 3446; DB 1; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.1e-178; Indels 0; Gaps 0;
Matches 660; Conservative 2; Mismatches 3;

QY	1	MSEKNFYITTPYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK	60
Db	15	MSEKNFYITTPYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK	74
QY	61	AREAGITPOAYDGMAGVVKELWQLLDISYDKFIETDDYHEKVAQVFERLLAODDIYL	120
Db	75	AREAGITPOAYDGMAGVVKELWQLLDISYDKFIETDDYHEKVAQVFERLLAODDIYL	134
QY	121	GEVSGWYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVEWVSESYFLRLSKYQDR	180
Db	135	GEVSGWYSVSDDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVEWVSESYFLRLSKYQDR	194
QY	181	LVEFFKAHPEFTTPDGRNLNEMLRNFIETDGLAVSRFTFTWGVVPNSPKHVVYVWIDA	240
Db	195	LVEFFKAHPEFTTPDGRNLNEMLRNFIETDGLAVSRFTFTWGVVPNSPKHVVYVWIDA	254
QY	241	LLNYATALGAQDEHGNFDFKFWNGTVFHWVGKDLIRFHSIYWPIILLMLDLVKLPRLIAH	300

Sequence 63402, A
Sequence 68671, A
Sequence 56601, A
Sequence 75322, A
Sequence 76244, A
Sequence 9544, Ap
Sequence 60052, A
Sequence 6961, A
Sequence 55682, A
Sequence 61147, A
Sequence 78246, A
Sequence 65466, A
Sequence 48489, A
Sequence 58468, A
Sequence 51105, A
Sequence 44568, A
Sequence 67599, A
Sequence 63291, A

Db 255 LNNYATAGYGAQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMLDLVKLPDLIAH 314
QY 301 GWFVKDGMKSKGNVVTPEMLVERYGLDPLRYILMRNLPVGSDGTFPTPDYVGRINYE 360
Db 315 GWFVKDGMKSKGNVVTPEMLVERYGLDPLRYILMRNLPVGSDGTFPTPDYVGRINYE 374
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
Db 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTHMEAVDPR 434
QY 421 ALBAVMTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
Db 435 ALBAVMTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETIAYIKQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETIAYIKQMEGNKPAV 554
QY 541 EKWNPDVELKLNKDEIKFDFKVEIRVAEVEKSKVEGSDKLLQPLDAGDGEDRQI 600
Db 555 EKWNPDVELKLNKDEIKFDFKVEIRVAEVEKSKVEGSDKLLQPLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 2

PCT-US02-09107B-74415
; Sequence 74415, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 74415
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
PCT-US02-09107B-74415

Query Match 82.9%; Score 2883.5; DB 1; Length 666;
Best Local Similarity 81.9%; Pred. No. 2.4e-148;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGKIQK 60
Db 1 MCKKPYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGHEVFLYGLDGHGKIQK 60
QY 61 ABEAGITPQAYVGMVAGVKELWQLLDISYDKPRTDDYHEKVAQVPERLLAODDIYL 120
Db 61 ABEAGITPQAYVGMVAGVKELWQLLDISYDKPRTDDYHEKVAQVPERLLAODDIYL 120
QY 121 GEYSGWYSVDEBFFTSQAEVFRDAGNVTGGTAPSGHEVWYSSESYFLRLSKYQDR 180

Db 121 GEYSGWYSVDEBFFTSQAEVFRDAGNVTGGTAPSGHEVWYSSESYFLRLSKYQDR 180
QY 181 LVPEFFKAHPFIITPDGRLNEMLNFIETPGLDIAVSRRTFTTGWGPVPSNPKHVVYVWIDA 240
Db 181 LVPAFFKEPFIITPDGRLNEMLNFIETPGLDIAVSRRTFTTGWGPVPSNPKHVVYVWIDA 240
QY 241 LLNATATAGYGAQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMLDLVKLPDLIAH 300
Db 241 LLNATATAGYGAQDEHGFDFKFWNGTVFHVGVKDLIRFHSIYWPILLMLDLVKLPDLIAH 300
QY 301 GWFVKDGMKSKGNVVTPEMLVERYGLDPLRYILMRNLPVGSDGTFPTPDYVGRINYE 360
Db 301 GWFVKDGMKSKGNVVTPEMLVERYGLDPLRYILMRNLPVGSDGTFPTPDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
Db 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDPR 420
QY 421 ALBAVMTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
Db 421 ALBAVMTLISRTNKYIDETAPWVLDKDEALRDQASVMSHWQASIRVVAHLIEPFFMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETIAYIKQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETIAYIKQMEGNKPAV 540
QY 541 --EKWNPDVELKLNKDEIKFDFKVEIRVAEVEKSKVEGSDKLLQPLDAGDGEDR 598
Db 540 SQEKWVPEEVALKSEKDVITFTFDVAEIRVAEVEKSKVEGSEKLLAFRVDAGDQDR 599
QY 599 QILSGIAKYYPNEQELVGVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPV 658
Db 600 QILSGIAKYYPNEQELVGVKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPV 659
QY 659 PNGSVIG 665
Db 660 PNGSVIG 666

RESULT 3

PCT-US02-09107B-72423
; Sequence 72423, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 72423
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Streptococcus mutans
PCT-US02-09107B-72423

Query Match 82.9%; Score 2883; DB 1; Length 669;
Best Local Similarity 82.4%; Pred. No. 2.6e-148;
Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;
QY 1 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGKIQK 59
Db 1 MTEKQPYITPTIYPSGKLHIGSAVTTIACDVLARYKRMNHDVFLYGLDGHGKIQK 60

```
QY 60 KAEAGITPOAYVDGMAGVKEMLQQLDISYDKFIRTTDDYHKKVAQVFERLLAQQDIY 119
Db 61 KSEAGITPOAYVDGMAGVKEMLQQLDISYDKFIRTTDDYHKKVAQVFERLLAQQDIY 120
QY 120 LGYSGWYSVDSDEFFTESQLAVERDEAGNVTGGIAPSGHEVWVSEESYFRLRSKYOD 179
Db 121 LGYSGWYSVDSDEFFTESQLEIFRDEDKVJGGIAPSGHEVWVSEESYFRLRSNYAD 180
QY 180 RLVEFFKAHPEFITPDGRNLEMLRNFIPEGLEDLAVSRITFTTGWVPVPSNPKHVYVWID 239
Db 181 RLVDFFHAHDFIOPGRMNEIKNFIPEGLEDLAVSRITFTTGWVPVPSNPKHVYVWID 240
QY 240 ALLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMLLDVCLPRLIA 299
Db 241 ALLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMLLDVCLPRLIA 300
QY 300 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINX 359
Db 301 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINX 360
QY 360 ELANDIGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 419
Db 361 ELANDIGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMET 479
Db 421 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMET 480
QY 480 SRAVLTOIGLEBVSSENLSLADFPADVTYVAKGTPIFPPLNMBEETIAYKEQM-- 539
Db 481 SNAIMEQLGAGQDLEGLAGLPTGKVIKGTPIFPPLNMBEETIAYKEQM-- 539
QY 540 V----EKEWNPDEVELKLNKDETFKFDKVEIRVAEVEKSVKVGSDKLLQRLDAGD 595
Db 540 IAGEEKEWNPDEVELKLNKDETFKFDKVEIRVAEVEKSVKVGSDKLLQRLDAGD 599
QY 596 EDQILSGIAKYYPNEQELVGKVKQIVANLKPMMKKYVSQGMILSAE--DGKLTLLTV 654
Db 600 EDQILSGIAKYYPNEQELVGKVKQIVANLKPMMKKYVSQGMILSAE--DGKLTLLTV 658
QY 655 DPAVNGSVIG 665
Db 659 DSSVNGSQIG 669
RESULT 4
PCT-US02-09107B-57809
; Sequence 57809, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57809
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Enterococcus faecium
PCT-US02-09107B-57809
```

```
Query Match 66.0%; Score 2295; DB 1; Length 669;
Best Local Similarity 65.2%; Pred. No. 1.4e-116; Indels 12; Gaps 9;
Matches 438; Conservative 93; Mismatches 129;
QY 1 MSEK-NFYITTFIYPPSGKLHIGSAYTTIACDLVARYKRLMGYDFVYLTGJDEHGKTIQQ 59
Db 1 MAEKSETFYITTFIYPPSGKLHIGSAYTTIACDLVARYKRLMGYDFVYLTGJDEHGKTIQQ 60
QY 60 KAEAGITPOAYVDGMAGVKEMLQQLDISYDKFIRTTDDYHKKVAQVFERLLAQQDIY 119
Db 61 KAEAGITPOAYVDGMAGVKEMLQQLDISYDKFIRTTDDYHKKVAQVFERLLAQQDIY 120
QY 120 LGYSGWYSVDSDEFFTESQLAVERDEAGNVTGGIAPSGHEVWVSEESYFRLRSKYOD 179
Db 121 LGYSGWYSVDSDEFFTESQLEIFRDEDKVJGGIAPSGHEVWVSEESYFRLRSNYAD 180
QY 180 RLVEFFKAHPEFITPDGRNLEMLRNFIPEGLEDLAVSRITFTTGWVPVPSNPKHVYVWID 239
Db 181 RLVEFFKAHPEFITPDGRNLEMLRNFIPEGLEDLAVSRITFTTGWVPVPSNPKHVYVWID 240
QY 240 ALLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMLLDVCLPRLIA 299
Db 241 ALLNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMLLDVCLPRLIA 299
QY 300 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINX 359
Db 300 HGFVMDKGMKSKGNVVPPEMLVERYGLDPLRYILMRNLPGVSDGTFPTPEDYVGRINX 359
QY 360 ELANDIGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 419
Db 360 ELANDIGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYP 419
QY 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMET 479
Db 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMET 479
QY 480 SRAVLTOIGLEBVSSENLSLADFPADVTYVAKGTPIFPPLNMBEETIAYKEQM-- 535
Db 480 SRAVLTOIGLEBVSSENLSLADFPADVTYVAKGTPIFPPLNMBEETIAYKEQM-- 539
QY 536 -NKPAVEKEWNPDEVELKLNKDETFKFDKVEIRVAEVEKSVKVGSDKLLQRLDAG 593
Db 540 TNEDEV--KWDPEETELVSTKEQIKFDVFEKVELKVAEIVNCKVEGADKLLQRLDAG 597
QY 594 DGEDRQILSGIAKYYPNEQELVGKVKQIVANLKPMMKKYVSQGMILSAE--HDGKLTLL 652
Db 598 DGEDRQILSGIAKYYPNEQELVGKVKQIVANLKPMMKKYVSQGMILSAE--HDGKLTLL 656
QY 653 TVDPAVNGSVI 664
Db 657 EAPKSPNGSEI 668
RESULT 5
PCT-US02-09107B-42475
; Sequence 42475, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
```

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 42475

; LENGTH: 677

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

PCT-US02-09107B-42475

Query Match

65.8%; Score 2288; DB 1; Length 677;

Best Local Similarity 63.9%; Pred. No. 3.4e-116;

Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPIYPSGKLHIGSAVTIACDVLARYKELMGYDVLYLTGLDEHGQKIQQ 59

DB 9 MSQKETFYITPIYPSGKLHIGNSYTIACDAMARYKELMGDFVLYLTGVDEHQKLEK 68

QY 60 KAEAGITPOAYVDGMAGVVKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQDDIY 119

DB 69 KAAELNVTPEYVDKMAADYQKLWKTLDISYDKPIRTTDDYHMAAVQOIFDLRVEQSDIY 128

QY 120 LGEYSGWYSVSDDEEFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSESYFLRLSKYQD 179

DB 129 LGEYEGWYSVSDDEEFTETQLAEVYRDDGNGVIGGKAPSGHEVELVKESYFFRMSKYAD 188

QY 180 RLVEFFKAHPEFITPQGRNLNEMLNPIEPGLEDLAVSRITFTTWGVPVPSNPKHVVYVWID 239

DB 189 RLVOYVEEHPFITQEPSRKNEMLNFIKPGLEDLAVSRITFTTWGVIPLKNDPKHVVYVWID 248

QY 240 ALNAYATAGYAQDEHGNDFKNGTVFHMVGKDLIRFHSIYWPILLMLLVKLPDLRLIA 299

DB 249 ALSNYITALGYGSEDDSLFQKYPANV-QVNGKEIVRFHTIYWPIMLMALDPLPKKVPFG 307

QY 300 HGVFMKDGKMSKGNVYVPEMLVERYGLDPLRYVYLMENLPVSGDGTPTPDYVGRINY 359

DB 308 HGWLLMKDGKMSKGNVYVPEMLVERYGLDALRYVYLLRAIPFGSDGVFTPDYVSRINY 367

QY 360 ELANDGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPFTHMEADVY 419

DB 368 DLANDGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPFTHMEADVY 427

QY 420 RALEAVWTISRTNKYIDETAPWLDKDEALRDQALASVNSHWQASIRVVAHLIEPFMMET 479

DB 428 TATAEITWLVSRANKYIDETAPWLVAKESKEKELESVMIHLAESURIVAILLQPMET 487

QY 480 SRVAVTQIGLE-EVSSLENSLADFPADVTVVAKGTPPIPRLMMEETIAYIKQEMGNKP 538

DB 488 PGKIFQLGLDPTMNMENIHFEFFPTDVTVTSKGTPIPRLEIETEVYIQQKMSQSES 547

QY 539 AVEKE--WNPDEVELKND-EIKFDFDKVEIRVAEVKEVSKVSGSKLQFRLDAGD 595

DB 548 ATBEDIKWPEETLVSTKEQIKYDDFDKVELKVAEVIDCKVKGADKLLQFRLDAGD 607

QY 596 EDQILSGIAKYYPNEQELVKKVQIVANLKPRKMKKYVSGMILSAEH-DGKLTLLTV 654

DB 608 NHRQILSGIAEFPDPAALIGKVVIVANLKPRK-MRGOISQGMILSAESPGLQIVEA 666

QY 655 DPAVNGSVI 664

DB 667 PKEMPNGAGI 676

RESULT 6

PCT-US02-09107B-60385

; Sequence 60385, Application PC/TUS0209107B

; GENERAL INFORMATION:

; APPLICANT: Elitra Pharmaceuticals Inc.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034VPC

; CURRENT APPLICATION NUMBER: PCT/US02/09107B

; CURRENT FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 09/948,993

; PRIOR FILING DATE: 2001-09-06

; PRIOR APPLICATION NUMBER: 60/342,923

; PRIOR FILING DATE: 2001-10-25

; PRIOR APPLICATION NUMBER: 10/072,851

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: 60/362,699

; PRIOR FILING DATE: 2002-03-06

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60385

; LENGTH: 664

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

PCT-US02-09107B-60385

Query Match

60.7%; Score 2113; DB 1; Length 664;

Best Local Similarity 60.7%; Pred. No. 9.1e-107;

Matches 407; Conservative 101; Mismatches 144; Indels 18; Gaps 8;

QY 3 EKN-FYITPIYPSGKLHIGSAVTIACDVLARYKELMGYDVLYLTGLDEHGQKIQKA 61

DB 4 EKNTFYITPIYPSGKAHIGHAYTVAGDAMARYKELXGYDVLYLTGTDDEHQKIQKA 63

QY 62 BEAGITPOAYVDGMAGVVKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQDDIYLG 121

DB 64 KERGISQEVYDIAEGFQELWKLKLEISNTDFIRTTQDRHKTSEKIFEQLEQSDIYLG 123

QY 122 EYSGWYSVSDDEEFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSESYFLRLSKYQDRL 181

DB 124 EYEGWYSVSDDEEFTETQLAEVYRDDGNGVIGGKAPSGHEVELVKESYFFRMSKYADRL 183

QY 182 VSEFFKAHPEFITPQGRNLNEMLNPIEPGLEDLAVSRITFTTWGVPVPSNPKHVVYVWIDAL 241

DB 184 VEYNSHPFELPESRKNEMLNFIKPGLEDLAVSRITFTTWGVIPLKNDPKHVVYVWIDAL 243

QY 242 LNYATAGYAQDEHGNDFKNGTVFHMVGKDLIRFHSIYWPILLMLLVKLPDLRLIAHG 301

DB 244 SNYITALGYTNDTKFQKYPADV-QVNGKEIVRFHTIYWPIMLMALDPLPKVFGHG 302

QY 302 WFMKDGKMSKGNVYVPEMLVERYGLDPLRYVYLMENLPVSGDGTPTPDYVGRINYEL 361

DB 303 WILMKDGKMSKGNVYVPEMLRIDRYGLDALRYVYLLREVPPFGSDGLFTPEDFVDRVNFOL 362

QY 362 ANDGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPFTHMEADVYDRA 421

DB 363 ANDGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPFTHMEADVYDRA 422

QY 422 LEAVWTISRTNKYIDETAPWLDKDEALRDQALASVNSHWQASIRVVAHLIEPFMMETSR 481

DB 423 LNQLWSLISRTNKYIDETAPWLVAKESKEKELESVMIHLAESURIVAILLQFLTRTPG 482

QY 482 AVTQIGLEEVSSLENI-----SLADFPADVTVVAKGTPPIPRLMMEETIAYIKQEM 534

DB 483 EIFLQLGLQE-----ENLKKWDSIYGYGEIPAGTVVKKGTPIPRLEDAEVETIYIQEMK 538

QY 535 GNPFAVEKWNPDVELKLNKDKIPEDPKVIRVAEVKEVSKVSGSKLQFRLDAGD 594

DB 539 GSAPAPAE--VAEVE-ALETPOIGEDFDKIDIRVAEVKQVDKVKKADKLLCFQDLG 595

QY 595 GEDRQLSGIAKYYPNEQELVKKVQIVANLKPRKMKKYVSGMILSAEHDKLTLLTV 654

DB 596 GKLRQLSGIAEFPDPAALIGKVVIVANLKPRK-MRGLMSEGMILSGEKDGKLSVIEA 653

QY 655 DPAVNGSVI 664

DB 654 SSALPNGAKV 663

RESULT 7

US-60-622-712-354

; Sequence 354, Application US/60622712

; GENERAL INFORMATION:

; APPLICANT: Klaenhammer, Todd

; APPLICANT: Russell, William


```

RESULT 9
PCT-US02-09107B-46123
; Sequence 46123, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 46123
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US02-09107B-46123

Query Match      55.6%; Score 1935.5; DB 1; Length 669;
Best Local Similarity 57.0%; Pred. No. 3.5e-97;
Matches 383; Conservative 107; Mismatches 161; Indels 21; Gaps 12;

Qy 1 MSEKN--FYITTPYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQIKI 58
Db 10 MTEENKSFYITTPYPSGKLHIGAYTTVAGDAMARYKCMQGVNVHYLTGTDEHGQIKI 69

Qy 59 QKABEAGITPOAYVDMGAVGKELWQLDIDSYDKFTRTDDYHEKVAQVFRLLAQQDI 118
Db 70 KKAELNVTPOAYVDMINAVIGELWEMDIDSYDDFIRTTEDRHKDVVEKIFQLVDQGI 129

Qy 119 YLGEYSGWVSDEEFFTSQLAEVRDEAGNVGTGIAP-SGHEVEWVSEESYFRLSKY 177
Db 130 YLDEYSGWVSQDETYTEHQLVDPIH-EGDKVVGKSPDSGSHVELVREESYFRMGKY 188

Qy 178 QDLVFEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRITFTTGWGVPVSPNPKHVYVW 237
Db 189 VDRLLKFYEDNPHFIQPSRKNEMINNFIKPGLDLAVSRITFTTGWGVPVSPNPKHVYVW 248

Qy 238 IDALLNATAGYAOQDEHGNDFKFWNGTVFHMVGKDILRFHSIYWPILLMLDVLDPRL 297
Db 249 VDALSNTITAGTGTEENEMKFWPADV-HLVGKEIVRFHTIYWPILLMALDPLPKKV 307

Qy 298 IAHGFWFMKDGKMSKGNVYVPEMLVERYGDLPLRYILMRNLPGVSGDTFTTPEYVGR 357
Db 308 FAHGWMILMDGKMSKGNVDPVILIDRYGLDALRYLLREVPPGSDGVFTTPEGFVER 367

Qy 358 NYELANDLGNLNRVTSYMNKYFDGQIPAYVGVTEPFDHVLAEVAKSIADPHTHMEAVD 417
Db 368 NFPLANDLGNLNRVTSYMNKYFDGQIPAYVGVTEPFDHVLAEVAKSIADPHTHMEAVD 427

Qy 418 YPRAEAVNTLISRTNKYIDETAPWLDKDEALRDLQASVMHQAOSIRVVAHLTEPFMM 477
Db 428 FSVALSISIQWLSRTNKYIDETAPWLDKDEALRDLQASVMHQAOSIRVVAHLTEPFMM 487

Qy 478 ETSRAVLTOIGL--EVSSLENLS-LADPPADVTVVAKGTPIPPRLNMBEETAYIKEQME 534
Db 488 VAPSRKFAQLGLTDEAHTSWGSLSTIGCIPAG-TKYVEKGQPIFPRLMDVVEYAYIKEQMK 546

Qy 535 GKNPAYE--KENPNPDEVELKLNKDEKDFDKVIRVAEVEKSVKSGSDKLLQFLRDA 592
Db 547 ASAPKVEEKEEPP-----KAEETIDDPFKVELRVAEVLASPEVKKADKLLQIDL 599

Qy 593 GGDGDRQILSGIAKYPNQELGVKKVQIVANLKPDKMMKKYVSGMILSAEHDGKLTLL 652

```

```

Db 600 GT-EKRQVVSIGIAKPY-SPEDLKGKVKVCTNLPKVK-LRGELSQGMILAGEENGULSLA 656
Qy 653 TVDPAPVNGSVI 664
Db 657 SIDQNIENGTKI 668

RESULT 10
PCT-US02-09107B-43933
; Sequence 43933, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 43933
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
PCT-US02-09107B-43933

```

```

Query Match      55.5%; Score 1931; DB 1; Length 657;
Best Local Similarity 56.3%; Pred. No. 6e-97;
Matches 377; Conservative 110; Mismatches 163; Indels 20; Gaps 11;

Qy 1 MSEKNFYITTPYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQIQK 60
Db 1 MAKETFYITTPYPSGKLHIGAYSTVAGDVIARYKRMQGYDVRVLTGTDEHGQIQK 60

Qy 61 ABEAGITPOAYVDMGAVGKELWQLDIDSYDKFTRTDDYHEKVAQVFRLLAQQDIYL 120
Db 61 AQAKGTEIYLDDEMTAGIKQLWAKLEISNDDEIRTEERHKKHVVQVPERLLKQGDYIL 120

Qy 121 GEYSGWVSDEEFFTSQLAEVRDEAGNVGTGIAP-SGHEVEWVSEESYFRLSKYQD 179
Db 121 GEYSGWVSDEEFFTSQLAEVRDEAGNVGTGIAP-SGHEVEWVSEESYFRLSKYQD 179

Qy 180 RLVEFFKAHPFITPDGRNLNMLNFIPEGLEDLAVSRITFTTGWGVPVSPNPKHVYVW 239
Db 180 RLLEFYDQNDPFIQPSRKNEMINNFIKPGLDLAVSRITFTTGWGVPVSPNPKHVYVW 239

Qy 240 ALLNATAGYAOQDEHGNDFKFWNGTVFHMVGKDILRFHSIYWPILLMLDVLDPRLIA 299
Db 240 ALVNYISALGYLSDDESLEFNKYWPADI-HLMAKEIVRFHSIYWPILLMALDPLPKKVFA 298

Qy 300 HGFWFMKDGKMSKGNVYVPEMLVERYGDLPLRYILMRNLPGVSGDTFTTPEYVGRIN 359
Db 299 HGWMILMDGKMSKGNVDPVILIDRYGLDALRYLLREVPPGSDGVFTTPEGFVERTNF 358

Qy 360 ELANDLGNLNRVTSYMNKYFDGQIPAYVGVTEPFDHVLAEVAKSIADPHTHMEAVDP 419
Db 359 DLANDLGNLNRVTSYMNKYFDGQIPAYVGVTEPFDHVLAEVAKSIADPHTHMEAVDP 418

Qy 420 RALEAVNTLISRTNKYIDETAPWLDKDEALRDLQASVMHQAOSIRVVAHLTEPFMMET 479
Db 419 VALSTVWKFSRTNKYIDETAPWLDKDEALRDLQASVMHQAOSIRVVAHLTEPFMMET 478

Qy 480 SRAVLTOIGLE-----EVSSLENLS-LADPPADVTVVAKGTPIPPRLNMBEETAYIKEQME 534
Db 479 PKSIFEQLNINNPQFMEFSLEQYGV--LTESIMVTGQPKPIFPRLDSEBIAIYKESMQ 536

```

QY 535 GNKPAVEKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKSGDKLQFRLDAGD 594
DB 537 --PPATEEE--KSEIP--SKPQIDIKDFDKVBIKAATIIDAHEVKSJLAKIQVDL-D 588
QY 595 GEDRQILSGIAKYPNEQELVGGKQVIVANLKPCKKKKVVSGMILSAEHGDKLTLTV 654
DB 589 SEQRQIVSGIAKPY--TPDDIIGKKVAVVTNLKPAKLMGQ--KSEGMLSAEKDGLTVLSL 646
QY 655 DPAVNGSVI 664
DB 647 PSAIPNGAVI 656
RESULT 11
PCT-US02-09107B-71042
; Sequence 71042, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71042
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
PCT-US02-09107B-71042

Query Match 55.3%; Score 1922.5; DB 1; Length 656;
Best Local Similarity 55.9%; Pred. No. 1.7e-96;
Matches 375; Conservative 113; Mismatches 160; Indels 23; Gaps 11;
QY 1 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKELMGYDVLYLTGLDEHGKIQOK 60
DB 1 MAKDTFYITPIYPSGKLHIGSAITTVAGDVIARYKMQGYDVRYLTGTDEHGKIQOK 60
QY 61 ABEAGITPOAYVDGMAVGKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
DB 61 AQAKGTELEVLDWISGINKLWSKLEISNDDFIRTEERHKQVVEKVERELLKQGDYL 120
QY 121 GEYSGWYSVDEBFFFTESQLAE-VFRDEAGNVTTGGIAP--SGHEVWVSSESYFRLSKYQ 178
DB 121 GEYEGWYSPDETYYTESQLDVPVY--ENGKIVGKSPDSGHEVELKEESYFFNINKYT 178
QY 179 DRLVEFFKAEFEFTPDGRINLEMLNFTPEGLDLAVSRRTFTTGWGPVPSNPKHVVYWI 238
DB 179 DRLLEFYDENDDFTQPPSRKKNEMNFIKPGLEDLAVSRFTSFDGVRVPSNPKHVVYWI 238
QY 239 DALLNATYALGAQOEHGDFKFWNGTVFHMVGKDILAFHSIYWPILMLMDVLPDLRI 298
DB 239 DALVNYISSGLYSDDETLENKYWPADI--HLMXKEIVRFHSIIPILMALDLPKXVF 297
QY 299 AHGFWFKDGMKSKSGKNVVPPEMLVERYGDPDLYLMLNLPVSGDGTFTPEYVGRIN 358
DB 298 AHGWTLMKDGKMSKSGKNVDPNVLIDRYGLDARYLMLRELPFGSDGVFTPEAFVERTN 357
QY 359 YELANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEFDHVLAEVAKESIADEPHTHEAVDY 418
DB 358 YDLANDLGNLNRVTSMINKYFHGELPAYQGPKEHDEKMEAMALETVKSFNDNMESLQF 417

QY 419 PRALAEVWTLISRTNKYIDETAPVLDKDEALRDQALASVMHQASIRVVAHLIEPFMME 478
DB 418 SVALSTVWKFISRTNKYIDETQPWVLAKDENQREMLGNVVAHLIVENIRFAIILQOFLTH 477
QY 479 TSRAVLTIQIG-----LEEVSLENLSLADFPADVTVVAKGTPPIFPRLNMESEIAYIKEM 533
DB 478 APREIFKQNLNNPDLHOLDLSIQVGM--LSEAITVTEKPTPIFPRLDTEAEIAYIKESM 535
QY 534 EGNKPAVEKWNPDVELKLNKDEIKFDPKVEIRVAEVEKSVKSGDKLQFRLDAG 593
DB 536 QPPK--SIKQSDPE-----GREQIDIKDFDKVBIKAATIIDAENVKSEKLLKIKVEL- 586
QY 594 DGEDRQILSGIAKYPNEQELVGGKQVIVANLKPCKKKKVVSGMILSAEHGDKLTLTV 653
DB 587 DNEQRQIVSGIAKPY--RPEDIIGKKVAVVTNLKPAKLMGQ--KSEGMLSAEKDGLTVLSL 644
QY 654 VDPAPVNGSVI 664
DB 645 LPSAIPNGAVI 655
RESULT 12
PCT-US04-37204-4240
; Sequence 4240, Application PC/TUS0437204
; GENERAL INFORMATION:
; APPLICANT: NOVOZYMES BIOTECH, INC.
; APPLICANT: NOVOZYMES A/S
; TITLE OF INVENTION: Bacillus Licheniformis Chromosome
; FILE REFERENCE: 10588.204-WO
; CURRENT APPLICATION NUMBER: PCT/US04/37204
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: 60/535,988
; PRIOR FILING DATE: 2004-01-09
; PRIOR APPLICATION NUMBER: 60/561,059
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 60/572,403
; PRIOR FILING DATE: 2004-05-19
; NUMBER OF SEQ ID NOS: 8395
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4240
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
PCT-US04-37204-4240

Query Match 54.9%; Score 1911; DB 1; Length 662;
Best Local Similarity 56.4%; Pred. No. 7.2e-96;
Matches 375; Conservative 107; Mismatches 173; Indels 10; Gaps 8;
QY 3 EKNFYITPIYPSGKLHIGSAITTIACDVLARYKELMGYDVLYLTGLDEHGKIQOKA 61
DB 4 EKNTFYITPIYPSGKLHIGSAITTVAGDAMARYKRLGFDVRYLTGTDEHGKIQOTA 63
QY 62 BEAGITPOAYVDGMAVGKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 121
DB 64 EKENITPQBLVDRAEDIQOLWKLLDISNDDFIRTEERHKVIEKVFQKLLDNGDIYLD 123
QY 122 EYSGWYSVDEBFFFTESQLAEVFRDEAGNVTTGGIAP--SGHEVWVSSESYFRLSKYQOR 180
DB 124 EYEGWYSIPDETYYTETQLVDVERNEKGEVIGGKSPDSGHPVELIKEESYFFRMGYADR 183
QY 181 LVEFFKAEFEFTPDGRINLEMLNFTPEGLDLAVSRRTFTTGWGPVPSNPKHVVYWI 240
DB 184 LLAFYENPEFTQPPSRKKNEMNFIKPGLEDLAVSRFTSFDGKIVGPNKPVYWI 243
QY 241 LLNATYALGAQOEHGDFKFWNGTVFHMVGKDILAFHSIYWPILMLMDVLPDLRI 300
DB 244 LFNYITALGENTENDENYQKYPADV-HLVGKEIVRFHSIIPILMALDLPKXVF 302
QY 301 GWFWMKDGKMSKSGKNVVPPEMLVERYGDPDLYLMLNLPVSGDGTFTPEYVGRINYE 360
DB 303 GWTLMKDGKMSKSGKNVDPNVLIDRYGLDARYLMLRELPFGSDGVFTPEGEVERINYD 362

QY 361 LANDGLNLRNTSMINKYFGQIPAYVEGVTFBDFHVLAEVAEKSIADPHTHEAVDYPR 420
Db 363 LANDGLNLRNTSMINKYFGQIPAYVEGVTFBDFHVLAEVAEKSIADPHTHEAVDYPR 422
QY 421 ALBAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMHQAQSIKQVVAHLIPEPFMMETS 480
Db 423 ALSSLMQFVSRNTNKYIDETAPWVLDKDEALRDOLASVMHQAQSIKQVVAHLIPEPFMMETS 482
QY 481 RAVTQIGLEEVSLNLSLADPADVTYVAKGTIPFRLNMEETIAYI----KEQMEGN 536
Db 483 EKIPAQGLVTDASLKTWDSQSGQLKSVTVKQGEPLFRLEAEDEVAYIKMQQTAP- 541
QY 540 VEKWNDEVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQFLDAGDGEDRQ 599
Db 542 --KEEPQKEEKAHERLPEITIDDFMSTELRAVEIHAEPVKKADRLKQLDLG-PEKEQ 598
QY 600 ILSGIAKYYPNEQELVGVKQVQIVANLKPROMKKYVSGQMILSAEHDGKLTLLTVDPAVP 659
Db 599 VWSGIAKHYKPE-ELVGRKVICVTLNKPVK-LRGEISQGMILAGEDNGVLSAAVDSSLA 656
QY 660 NGSVI 664
Db 657 NGTRI 661

RESULT 13

PCT-US02-09107B-52642
; Sequence 52642, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52642
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Clostridium botulinum
PCT-US02-09107B-52642

Query Match 48.2%; Score 1678; DB 1; Length 674;
Best Local Similarity 48.7%; Pred. No. 2.8e-83;
Matches 326; Conservative 136; Mismatches 176; Indels 32; Gaps 10;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFVLTGLDGHGKIQOK 60
Db 30 MSRGTYVITPIYPSAKLHIGNYTTVAADTLARFKLTGYDVNMFVLTGTDBHGKIQRL 89
QY 61 ABEAGTTPQAYVDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIY 120
Db 90 ABAKVTTPKYVDIYVAGIKELWQNNIDYKFIKRTIDYHVELVQKFKLYDYGDIYK 149
QY 121 GEYSGWYSDEEFFTESQALAEVFRDEAGNVTGGIAPSGHEVSEYSEFYRLSKYQDR 180
Db 150 SEYSGWYCTPCSEFWTETQLED-----GKPCDCRSVEKAKEEAYFPKSKYAPK 199
QY 181 LVEFFKAHPEFTIPDGLNLEMLNFIETGLEDLAYSRTFTTGVGVPSNPKHVVVWIDA 240
Db 200 LLEYTESHEPEFQPEBRKKNMNNFUKPLQDLQCVSRSTFSDGWIPIVTFDEKRVIVYVWDA 259
QY 241 LLNATYALGYAQDEHGNFDFKWNVTFFHVMGVKDLIRFHSIYWPILLMLLDVKKLPRLIAH 300

Db 260 LANVITAGVSTEDKDELYKKTWPADV-HLVKQDLRPHHTIYWPIMLMALDIPLPKQVFGH 318
QY 301 GWFMKDGKMSKSGNVVYPEMLVERYGDLDFRYLLMRNLPVSGDGTFTTPEYVGRINYE 360
Db 319 GWWLVDDGKMSKSGNVVDPVWLVNHFCTDPVRYLLREIIPFGADGIFNNEIFIKKINS 378
QY 361 LANDGLNLRNTSMINKYFGQIPAYVEGVTFBDFHVLAEVAEKSIADPHTHEAVDYPR 420
Db 379 LANDGLNLRNTSMINKYFGQIPAYVEGVTFBDFHVLAEVAEKSIADPHTHEAVDYPR 437
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDOLASVMHQAQSIKQVVAHLIPEPFMMETS 480
Db 438 ALSEIWTIFGRANKYIDETTPWILAKDESKKERLGTVLVNLASLRISVCISAFIPETS 497
QY 481 RAVTQIGLEEVSLNLSLADPADVTYVAKGTIPFRLNMEETIAYI----KEQMEGN 536
Db 498 IKINEQIN-ADVLNWSLSKFNKGTGKVKGEAIFPRIDVEKKIBELNKLKEEQLKQN 556
QY 537 KPAVEKSWNPDEVELKLNKDEIKFEDFDKVEIRVAEVEKSVKVEGSKLLQFLDAGDGE 596
Db 557 KAK-----NIAPIKEIITIDFEKIDLRVAKVLECEPIKGAKKLLKLVSLG-GE 605
QY 597 DRQILSGIAKYYPNEQELVGVKQVQIVANLKPROMKKYVSGQMILSA--EHDGKLTLLTV 654
Db 606 ERQVVSIGIAKYKPE-DLICKKVLVNLKPVK-LRGEISQGMILAASTEDDSELSFVASI 663
QY 655 DPAVNGSVI 664
Db 664 DGDIEGNIV 673

RESULT 14

PCT-US02-09107B-51647
; Sequence 51647, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51647
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
PCT-US02-09107B-51647

Query Match 47.2%; Score 1640.5; DB 1; Length 644;
Best Local Similarity 48.3%; Pred. No. 2.7e-81;
Matches 322; Conservative 125; Mismatches 193; Indels 27; Gaps 10;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDYFVLTGLDGHGKIQOK 59
Db 1 MSEKFFYITPIYPSAKLHIGNYTTVASDALVFRFKLQGYDAFMLTCTDBHGKIQOR 60
QY 60 KAEAGTTPQAYVDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODDIY 119
Db 61 IAEKGTTPKYVDIYVAGIKELWQNNIDYKFIKRTIDYHVELVQKFKLYDYGDIY 120
QY 120 LGEYSGWYSDEEFFTESQALAEVFRDEAGNVTGGIAPSGHEVSEYSEFYRLSKYQDR 179

Db	121	KSAYEGHYCTPCSFWTETQL-----VDGKCPDGRPVVEKTEKAEYFFKMSKYAD	170
Qy	180	RLVEFFKAHPEFTTPDGRNLEMLNFIEPGLEDLAVSRITFTWGVVPVPSNPKHVYVWID	239
Db	171	RLIKYIEDHDFDIPQPSRKNEMLNFKGLQDLCSRSFDMGIPITTEDEKHVIYVWID	230
Qy	240	ALLNATYALGYAODEHGNFDPKFWNGTVFHMVGKDIILRPHSIYWPILLMMLDVKLPDRLLA	299
Db	231	ALSNYITALGYGSDNDELNKKFWPADL-HLVGKDIIRFHTIYWPIMLMALDPLPKQVFG	289
Qy	300	HGMFVKDGMKSKSGNVVPEMLVERYGLDPLRYLYLMRNLPGVSDGTFTPEDYVGRINY	359
Db	290	HGMLLVDDGKMSKSGNVVDPVVLNFEFGTDPVRYLLHLEIFPGSDGLFNEIFIKKINS	349
Qy	360	ELANDLGNLNRVTSMINKYFDQIPAYVEGVTEFPHVLAEVAKSIAOPHTHMEAVDYP	419
Db	350	DLANDLGNLVSRTAAEMIKYFDGSIQPPVD-KSEIDNELIDMAISPEKLDEDIKKLKP	408
Qy	420	RALEAVWTLISRTNKYIDETAPWLDKDEALRQDLASVMSHWQASIRVVAHLIEPFMMET	479
Db	409	EALSHINDLKRANKYIDETFPWLAKDENKARLGTVLNLVESLRFVATTLTPFLPET	468
Qy	480	SRAVLTOIGLEBVSSLENLSLADFPADVTVVAKGTPIPPLNMBEIEIAYIKQMEGNKPA	539
Db	469	GEKIKTQINT-ELDTWESLSAFDGTAGTKVSKGEVIFPRIDVDKKEELNKLKEEQKKA	527
Qy	540	VEKEWNPDEVELKLNKDEIKFEDPKVEISVAEVKESKVEGSDKLIQFRLDAGDGEDRO	599
Db	528	TRK-----MQPLKPEISIDVDKDLRVRVLECEPVKSKKLLKLVKELG-GEERQ	578
Qy	600	ILSGIAKYVNEQELVGKKYQIVANLKPRKMKKYVSGQMIL--SAFHGDKGLTLLVDPA	657
Db	579	VLSGISQFYKPE-DLIGKKVVLVANLAPAKLMGQ-LSQGMILAVATDDDSKLYTLDIPED	636
Qy	658	VPNGSVI 664	
Db	637	IPTGSIV 643	

RESULT 15

PCT-US02-09107B-53102
 ; Sequence 53102, Application PC/TUS0209107B
 ; GENERAL INFORMATION:
 ; APPLICANT: Elittra Pharmaceuticals Inc.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034VPC
 ; CURRENT APPLICATION NUMBER: PCT/US02/09107B
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 09/948,993
 ; PRIOR FILING DATE: 2001-09-06
 ; PRIOR APPLICATION NUMBER: 60/342,923
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: 10/072,851
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: 60/362,699
 ; PRIOR FILING DATE: 2002-03-06
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 53102
 ; LENGTH: 645
 ; TYPE: PRT
 ; ORGANISM: Clostridium difficile
 PCT-US02-09107B-53102

Query Match 45.2%; Score 1572; DB 1; Length 645;
 Best Local Similarity 48.0%; Pred. No. 1.4e-77;
 Matches 319; Conservative 113; Mismatches 205; Indels 28; Gaps 11;

Qy	1	MSEKNFYITPIYPSGKLHIGSAITACDVLYARKLMGYDYFYLTLGLDEHGKIQOK	60
Db	1	MSKPSFYVITPIYPSGGLHIGHTYSTVAADTIAEFKFCGVDVKFLAGTDEHGEKIQK	60

Search completed: December 17, 2004, 22:12:06
 Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:41:55 ; Search time 158 Seconds

(without alignments)
1509.842 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479

Sequence: 1 MSEKNFYITPIYVSGKLH.....DGKLLITVDPAVNGSVIG 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3479	100.0	665	2 AAW27660	Streptococcus pneumoniae
2	3446	99.1	665	6 ABU01176	Streptococcus pneumoniae
3	3446	99.1	679	4 AAU37825	Streptococcus pneumoniae
4	3446	99.1	679	6 ABU45969	Streptococcus pneumoniae
5	3429	98.6	679	8 ADK46236	Streptococcus pneumoniae
6	2893.5	82.9	666	5 ABP26526	Streptococcus pneumoniae
7	2893.5	82.9	666	6 ABU46491	Streptococcus pneumoniae
8	2883	82.9	669	6 ABU44499	Streptococcus pneumoniae
9	2732.5	78.5	665	5 ABP30434	Streptococcus pneumoniae
10	2732.5	78.5	682	5 ABP26525	Streptococcus pneumoniae
11	2392.5	68.8	662	5 ABBS4105	Streptococcus pneumoniae
12	2295	66.0	669	6 ABU29885	Streptococcus pneumoniae
13	2231.5	65.9	674	7 ADC96574	Streptococcus pneumoniae
14	2288	65.8	677	4 AAU35070	Streptococcus pneumoniae
15	2288	65.8	677	6 ABU44551	Streptococcus pneumoniae
16	2284.5	65.7	666	4 AAU33437	Streptococcus pneumoniae
17	2154	61.9	680	6 ADH08140	Streptococcus pneumoniae
18	2113	60.7	664	6 ABBS0126	Streptococcus pneumoniae
19	2113	60.7	664	6 ABU32461	Streptococcus pneumoniae
20	2040	58.6	652	6 ADH08138	Streptococcus pneumoniae
21	1949	56.0	659	6 ABU33600	Streptococcus pneumoniae
22	1939	55.7	657	6 ABRS9494	Streptococcus pneumoniae
23	1935.5	55.6	669	6 ABU18199	Streptococcus pneumoniae
24	1934	55.6	657	4 AAU36542	Streptococcus pneumoniae
25	1931	55.5	657	2 AAW20051	Methionyl tRNA synthetase

ALIGNMENTS

RESULT 1

AAW27660

ID AAW27660 standard; protein; 665 AA.

XX AC AAW27660;

XX DT 28-APR-1998 (first entry)

XX DE Streptococcus pneumoniae methionyl tRNA synthetase.

XX KW Methionyl tRNA synthetase; metS polypeptide; antibacterial; vaccine; immune response; meningitis; bacterial infection.

XX OS Streptococcus pneumoniae.

XX PN WO9739012-A1.

XX PD 23-OCT-1997.

XX PF 18-APR-1997; 97WO-US006555.

XX PR 18-APR-1996; 96GB-00007999.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PI Lawlor EJ;

XX DR WPI; 1997-526388/48.

XX DR N-PSDB; AAT88098.

XX PT Polynucleotide encoding Streptococcus pneumoniae methionyl tRNA synthetase - useful to diagnose, treat or prevent bacterial infection, especially meningitis.

XX PS Claim 12; Page 36-38; 47pp; English.

XX CC The present sequence represents methionyl tRNA synthetase (metS) from Streptococcus pneumoniae. The polynucleotide encoding metS can be used for the recombinant production of metS, which can be used to treat conditions requiring metS activity. Sequences antisense to the polynucleotide can be used to control metS expression. MetS, or metS expressing vectors can be used to induce an immune response, i.e. an antibody (Ab) and/or T cell response, against S. pneumoniae to protect against infection, or to screen for antagonists or agonists of the polynucleotide encoding metS or metS's activity, i.e. antibacterials. The polynucleotide, e.g. an anti-metS Ab, can be used to treat conditions requiring metS inhibition, e.g. an S. pneumoniae infection, particularly

ABU16009 Protein e
AAU37420 Staphyloc
ABR5493 S. aureus
ABM71940 Staphyloc
ABU43118 Protein e
ABP38505 Staphyloc
AAU34400 Staphyloc
ADH08136 Alloioococ
AAG81883 S. epider
ABU23723 Protein e
ABU24718 Protein e
ABU25178 Protein e
ADH86149 Enterococ
ABU18430 Protein e
ABU26515 Protein e
ABU48823 Protein e
AAU39661 Helicobac
AAW20999 H. pylori
ABU30785 Protein e
AAW75910 Helicobac

26 1931 55.5 657 6 ABU16009
27 1930 55.5 657 4 AAU37420
28 1930 55.5 657 6 ABR5493
29 1930 55.5 657 6 ABM71940
30 1922.5 55.3 656 6 ABU43118
31 1922.5 55.3 660 5 ABP38505
32 1884 54.2 652 4 AAU34400
33 1856 53.3 605 6 ADH08136
34 1787.5 50.8 613 4 AAG81883
35 1678 48.2 674 6 ABU24718
36 1640.5 47.2 644 6 ABU23723
37 1572 45.2 645 6 ABU25178
38 1492.5 42.9 493 7 ADH86149
39 1383 39.8 670 6 ABU18430
40 1189.5 34.2 628 6 ABU26515
41 1153.5 33.2 521 6 ABU48823
42 1152.5 33.1 656 4 AAU39661
43 1152.5 33.1 662 2 AAW20999
44 1148.5 33.0 650 6 ABU30785
45 1137.5 32.7 648 2 AAW75910

CC meningitis. A fragment of the polynucleotide encoding metS can be used as a
 CC probe to isolate full length or related sequences, or diagnose, e.g. by
 CC polymerase chain reaction, infection stage and type, including mutation
 CC and polymorphism detection. Diagnosis may also be achieved by detecting
 CC metS gene overexpression, e.g. by immunosay. The Ab can be used to treat
 CC infection, isolate or identify metS expressing clones, purify metS and as
 CC an immunoassay reagent. More generally, the products can prevent adhesion
 CC of bacteria to wounds and in dwelling devices, block metS protein
 CC mediated invasion of mammalian cells and block the normal progression of
 CC infection
 XX
 SQ Sequence 665 AA;

Query Match 100.0%; Score 3479; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 2e-284;
 Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK 60
 DB 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK 60

QY 61 ABEAGITPCAYVDGKAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
 DB 61 ABEAGITPCAYVDGKAVGKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120

QY 121 GEYSGWYSDEEFFTESQLAEVFRDEAGNVTTGGIAPSGHEVVEVSESYFLRLSKYQDR 180
 DB 121 GEYSGWYSDEEFFTESQLAEVFRDEAGNVTTGGIAPSGHEVVEVSESYFLRLSKYQDR 180

QY 181 LVEFFKAHFEFTTPGRLNEMLNIEPCGLEDLAVSRFTFTGCVVPSPNKHVVYVWIDA 240
 DB 181 LVEFFKAHFEFTTPGRLNEMLNIEPCGLEDLAVSRFTFTGCVVPSPNKHVVYVWIDA 240

QY 241 LNNYATALGYAODEGNFQKFWNGVVFHMGKDILRFHSIYWPILLMLDKLPDLIAH 300
 DB 241 LNNYATALGYAODEGNFQKFWNGVVFHMGKDILRFHSIYWPILLMLDKLPDLIAH 300

QY 301 GWFVKDGRMSKSGNVVYPPEMLVRYGLDPLRYLYLMNLPVSGDGTTPEDYVGRINYE 360
 DB 301 GWFVKDGRMSKSGNVVYPPEMLVRYGLDPLRYLYLMNLPVSGDGTTPEDYVGRINYE 360

QY 361 LANDLGNLNRVSMINKYFDQIPAYVEGVTEPDHVLAEVAEKSIADPHTHMAVDYPR 420
 DB 361 LANDLGNLNRVSMINKYFDQIPAYVEGVTEPDHVLAEVAEKSIADPHTHMAVDYPR 420

QY 421 ALEAVNTLISRTNKYIDETAPVWLKDQALRDQLASVMSHWQASIRVVVAHLIEPFFMETS 480
 DB 421 ALEAVNTLISRTNKYIDETAPVWLKDQALRDQLASVMSHWQASIRVVVAHLIEPFFMETS 480

QY 481 RAVLTQIGLEEVSSLENISLADFPADVTVAAGTPIFRLNNEBEEIAYIKQMEGNKPAV 540
 DB 481 RAVLTQIGLEEVSSLENISLADFPADVTVAAGTPIFRLNNEBEEIAYIKQMEGNKPAV 540

QY 541 EKEMNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVESGSKLQRLDAGDGEDRQI 600
 DB 541 EKEMNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVESGSKLQRLDAGDGEDRQI 600

QY 601 LSGIAKYPNEBELGKVKQIVIANLKPRMKKKYVSQGMILSAHDGKLTLLTVDPAVFN 660
 DB 601 LSGIAKYPNEBELGKVKQIVIANLKPRMKKKYVSQGMILSAHDGKLTLLTVDPAVFN 660

QY 661 GSVIG 665
 DB 661 GSVIG 665

RESULT 2
 ABU01176
 ID ABU01176 standard; protein; 665 AA.
 XX
 AC ABU01176;
 XX
 DT 23-OCT-2003 (revised)

DT 11-FEB-2003 (first entry)
 XX S. pneumoniae type 4 strain protein from coding region #749.
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 XX antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX Streptococcus pneumoniae; type 4 strain.
 OS
 XX WO200277021-A2.
 PN 03-OCT-2002.
 XX 27-MAR-2002; 2002WO-IB002163.
 PF 27-MAR-2001; 2001GB-00007658.
 FR (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 PA Masignani V, Tettelin H, Frazer C;
 XX WPI; 2003-040579/03.
 DR N-PSDB; ABX06461.
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 PS Claim 1; SEQ ID NO 1498; 56pp; English.
 XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 665 AA;

Query Match 99.1%; Score 3446; DB 6; Length 665;
 Best Local Similarity 99.2%; Pred. No. 1.2e-281;
 Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK 60
 DB 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLTGLDEHGKIQOK 60

Db 495 RAVLTQLGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 554
Qy 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVABVKEVSKVSGSKLLQRLDAGDGEDROI 600
Db 555 EKEWNPDEVELKLNKDEIKFEDPKVEIRVABVKEVSKVSGSKLLQRLDAGDGEDROI 614
Qy 601 LSGIAKYYPNEQELVGGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVNP 660
Db 615 LSGIAKYYPNEQELVGGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVNP 674
Qy 661 GSVIG 665
Db 675 GSVIG 679

RESULT 4
ABU45969 standard; protein; 679 AA.
XX AC ABU45969;
XX 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #31496.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Streptococcus pneumoniae.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0343233P.
XX PR 08-FEB-2002; 2001US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Irawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX PS N-P8DB; ACA49839.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 73893; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 679 AA;

Query Match 99.1%; Score 3446; DB 6; Length 679;
Best Local Similarity 99.2%; Pred. No. 1.3e-281;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSEKNFYITPIYPSGKLHIGSAVTIACDVLARYKRLMGYDVFYLTGLDEHGKIQK 60
Db 15 MSEKNFYITPIYPSGKLHIGSAVTIACDVLARYKRLMGYDVFYLTGLDEHGKIQK 74
Qy 61 ABEAGITPQAYVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODIYL 120
Db 75 ABEAGITPQAYVDGMAGVVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAODIYL 134
Qy 121 GEYSGWYSVSDERFFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSEESYFLRLSKYQDR 180
Db 135 GEYSGWYSVSDERFFTESQLAEVFRDEAGNVTGGIAPSGHEVEMVSEESYFLRLSKYQDR 194
Qy 181 LVEFFKAHPEFITPDGRLEMLRNFIPEGLEDLAVSRITFTGWVPVSPKHYVYVWIDA 240
Db 195 LVEFFKAHPEFITPDGRLEMLRNFIPEGLEDLAVSRITFTGWVPVSPKHYVYVWIDA 254
Qy 241 LLNYATAGVAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSHIYWPILLMLLDVLPRLIAH 300
Db 255 LLNYATAGVAQDEHGNFDFKFWNGTVFHMVGKDIILRFHSHIYWPILLMLLDVLPRLIAH 314
Qy 301 GWFVMKDGKMSKSGKNVYVPEMLVERYGDPRLRYILMRNLVPGSGGTFTPDYVGRINYE 360
Db 315 GWFVMKDGKMSKSGKNVYVPEMLVERYGDPRLRYILMRNLVPGSGGTFTPDYVGRINYE 374
Qy 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDYPR 420
Db 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADPHTHMEAVDYPR 434
Qy 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVAHLIEPFMMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWLDKDEALRDQASVMSHWQASIRVAHLIEPFMMETS 494
Qy 481 RAVLTQIGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSADFPADVTVAAGTPIFFRLNMEESIAVIKEQMEGNKPAV 554
Qy 541 EKEWNPDEVELKLNKDEIKFEDPKVEIRVABVKEVSKVSGSKLLQRLDAGDGEDROI 600
Db 555 EKEWNPDEVELKLNKDEIKFEDPKVEIRVABVKEVSKVSGSKLLQRLDAGDGEDROI 614
Qy 601 LSGIAKYYPNEQELVGGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVNP 660
Db 615 LSGIAKYYPNEQELVGGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAVNP 674
Qy 661 GSVIG 665
Db 675 GSVIG 679
RESULT 5
ADK46236
ID ADK46236 standard; protein; 679 AA.

XX ADK46236;
AC 20-MAY-2004 (first entry)
DT Streptococcus pneumoniae protein, Seq ID No 2751.
DE Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
DE Streptococcus pneumoniae.
KW Streptococcus pneumoniae.
OS US6699703-B1.
XX 02-MAR-2004.
XX 26-MAY-2000; 2000US-00583110.
XX 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewheat CE;
PI WPI; 2004-212399/20.
XX N-PSDB; ADK43575.
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX Disclosure; SEQ ID NO 2751; 301pp; English.
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX Sequence 679 AA;
Query Match 98.6%; Score 3429; DB 8; Length 679;
Best Local Similarity 98.8%; Pred. No. 3.4e-280;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQK 60
DB 15 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQK 74
QY 61 AEFAGITPOAYVDMGAVGKELQDLIDSYDKIRTTDDYHEKVAQVFERLLAQDDIYL 120
DB 75 AEFAGITPOAYVDMGAVGKELQDLIDSYDKIRTTDDYHEKVAQVFERLLAQDDIYL 134
QY 121 GEYSGWYSVSDDEFFFTESQALAEYFRBAGNVTGGIAPSGHEVEWVSEESYFLRLSKYQDR 180
DB 135 GEYSGWYSVSDDEFFFTESQALAEYFRBAGNVTGGIAPSGHEVEWVSEESYFLRLSKYQDR 194
QY 181 LVBEFFKAHPFITPDGRLNEMLNFTPEGLDIAVSRITFTWGPVPSKPKHVVYWIIDA 240
DB 195 LVBEFFKAHPFITPDGRLNEMLNFTPEGLDIAVSRITFTWGPVPSKPKHVVYWIIDA 254
QY 241 LLNVATALGAQDEHGNFDFWNGTVPFHWGKIDILRFHSIYWPILLMLDLVKLPDRLIAH 300
DB 255 LLNVATALGAQDEHGNFDFWNGTVPFHWGKIDILRFHSIYWPILLMLDLVKLPDRLIAH 314
QY 301 GWFVKDKGMSKSGNVVYPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 360

DB 315 GWFVKDKGMSKSGNVVYPEMLVERYGLDPLRYLLMRNLPGVSDGTFTPEDYVGRINYE 374
QY 361 LANDLGNLLARTVSMINKYFDGQIPAYVEGVTDFDHVLAESAFAKSIADFHTEAVDYP 420
DB 375 LANDFGLNLLARTVSMINKYFDGQIPAYVEGVTDFDHVLAESAFAKSIADFHTEAVDYP 434
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFFMETS 480
DB 435 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFFMETS 494
QY 481 RAVITQGLBEVSSLENLSLADPPADVTVAAGTPIPPRLNMBEEIAYIYKEQEGNKP 540
DB 495 RAVITQGLBEVSSLENLSLADPPADVTVAAGTPIPPRLNMBEEIAYIYKEQEGNKP 554
QY 541 EKEWNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVKSGDKLQFRLDAGDGEDRQI 600
DB 555 EKEWNPDEVELKLNKDEIKFDFKVEIRVAEVEKSVKSGDKLQFRLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKVOIVANLKPRLKMMKYVSQGMILSAEHDGKLTLLTVDP 660
DB 615 LSGIAKYYPNEQELVGKVOIVANLKPRLKMMKYVSQGMILSAEHDGKLTLLTVDP 674
QY 661 GSVIG 665
DB 675 GSVIG 679
RESULT 6
ABP26526
ID ABP26526 standard; protein; 666 AA.
XX ABP26526;
AC ABP26526;
DT 02-JUL-2002 (first entry)
DE Streptococcus polypeptide SEQ ID NO 2228.
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
OS Streptococcus pyogenes.
XX WO200234771-A2.
XX 02-MAY-2002.
XX 29-OCT-2001; 2001WO-GB004789.
XX 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;
PI Tettelin H;
XX WPI; 2002-352536/38.
DB N-PSDB; ABN67157.
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX Claim 1; Page 3374; 4525pp; English.
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus (Streptococcus agalactiae) or group A streptococcus (GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 666 AA;

Query Match 82.9%; Score 2883.5; DB 5; Length 666;
 Best Local Similarity 81.9%; Pred. No. 3.9e-234;
 Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY 1 MSEKPYITPTIYPSGKLHIGSAVTTTACDVLARYKELMGVDYFYLTLDEHGQKIQK 60
 DB 1 MKKKPFYITPTIYPSGKLHIGSAVTTTACDVLARYKELMGHEVFYLTGLDEHGQKQTK 60

QY 61 ABEAGITPQAYVDGMAGVKELWQLLDDISYDKFIRTTDDYHEKVVAQVFERLLAQDDIYL 120
 DB 61 AKEAGITPQYVDNMAKVKALWQLLDDISYDFIRTTDDYHEVVAQVFERLLAQDDIYL 120

QY 121 GYSGWYSDEEFTTESQLAVFDEAGNVTGGTAPSGHEVWVSESYFLRLSKYQDR 180
 DB 121 GYSGWYSDEEFTTESQLKVFDEAGNVTGGTAPSGHEVWVSESYFLRLSKYQDR 180

QY 181 LVEFFKAHPEFTTPGRNLEMLNIEPGLDLAVSRITFTWGVDPVSPKHVVVWIDA 240
 DB 181 LVAFKERDFTQPGRNEMVKNIEPGLDLAVSRITFTWGVDPVSPKHVVVWIDA 240

QY 241 LNYATALGYAQDEHGNDFKFWGTVFHMVGKIDILRPHSIYWPILLMMDLVKLPDLIAH 300
 DB 241 LNYATALGYRQANHANFDFKFWGTVFHMVGKIDILRPHSIYWPILLMMDLPMFDRLIAH 300

QY 301 GFVWMDGKWSKGNVYVPEMLVRYGLDPLRYVLMENLPVSGDGTFTPDYVGRINYE 360
 DB 301 GFVWMDGKWSKGNVYVPEMLVRYGLDPLRYVLMENLPVSGDGTFTPDYVGRINYE 360

QY 361 LANDLGNLNRVTSVINKYFDGQIPAYVEGVTEFDPHVLAEVAEKSIADFTCHMEADVPR 420
 DB 361 LANDLGNLNRVTSVINKYFDGQIPAYVDNGTAFDADLSQLIDAQADYVHKHMEADVPR 420

QY 421 ALEAVWTLISRNKYIDETAPVWLDKALRDOLASVMSHWQASIRVVAHLIEPMMETS 480
 DB 421 ALEAVWTLISRNKYIDETAPVWLDKALRDOLASVMSHWQASIRVVAHLIEPMMETS 480

QY 481 RAVLTQIGLEEVSSLENISLADFPADVTVAKGTPIFPLNMEERIIAYIKQEMGNKPAV 540
 DB 481 AAIMQGLLEPVSJSLTALADFPANTKVAKGTPIFPLNMEERIIAYIKQEMGNKPAV 540

QY 541 --EKWNPDEVELKUNKEIDFEDFKVEIRVAEVEKVSKEVSGDKLLQFLRDAGDGDR 598
 DB 540 SOKEWNPDEVELKUNKEIDFEDFKVEIRVAEVEKVSKEVSGDKLLQFLRDAGDGDR 598

QY 599 QILSGIAKYPNEQELVGVKQVIVANLPRKMKKYYVSGQMLSAEHDKLTLTVDPAP 658
 DB 600 QILSGIAKYPNEQELVGVKQVIVANLPRKMKKYYVSGQMLSAEHDKLTLTVDPAP 658

QY 659 PNGSVIG 665
 DB 660 PNGSVIG 666

RESULT 7
 ABU46491
 ID ABU46491 standard; protein; 666 AA.
 XX
 AC ABU46491;

XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #32018.
 XX
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 KW Streptococcus pyogenes.
 XX
 OS WO200277183-A2.
 XX
 PN 03-OCT-2002.
 XX
 PD 21-MAR-2002; 2002WO-US009107.
 XX
 PF 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 PR
 PR 25-OCT-2001; 2001US-0342923P.
 PR
 PR 08-FEB-2002; 2002US-00072851.
 PR
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Chlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA50361.
 DR
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 74415; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: the sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 666 AA;

Query Match 82.9%; Score 2883.5; DB 5; Length 666;
 Best Local Similarity 81.9%; Pred. No. 3.9e-234;
 Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;

QY	1	MEKQFYITPTIYPYSGKLHIGSA	120	1	MMKQFYITPTIYPYSGKLHIGSA	120	61	AEZAGITPOAYVDGMAVGKSLW	120	61	AXEAGITQTYVDNNAKDVKAL	120	121	GRYSGWYSVDSBEFTTESQLAE	180	121	GRYSGWYSVDSBEFTTESQLAE	240	181	LYVEFFKAHPERTTPOGRNEMLR	240	181	LVAFFKERPDFQPOGRNEMWKP	300	241	LLNYATALGYAQDEHGNDFKWN	300	241	LLNYATALGYROAHNDFKWN	360	301	GHFVWKDGKMSKGNVYPEML	360	301	GHFVWKDGKMSKGNVYPEML	420	361	LANDLGNLNRVTSMINKYFDG	420	361	LANDLGNLNRVTSMINKYFDG	480	421	ALEAVWTLISRNKIYDITAPWL	480	421	ALEAVWTLIARNKIYDITAPWL	540	481	RAVLTIQIGLEEVSLSLENI	540	481	AAIMAQGLGPEVSDLSLALAD	598	541	--EKWNPDEVELKUNKDEIK	598	540	SOQEKWNPDEEVALKEKQVIT	599	599	QILSGIAKYYPNEQELVGKKQ	658	600	QILSGIANFYNEQELVGKKQ	659	659	PNGSVIG 665	660	PNGSIIG 666
----	---	-------------------------	-----	---	-------------------------	-----	----	------------------------	-----	----	-----------------------	-----	-----	------------------------	-----	-----	------------------------	-----	-----	-------------------------	-----	-----	------------------------	-----	-----	------------------------	-----	-----	----------------------	-----	-----	----------------------	-----	-----	----------------------	-----	-----	-----------------------	-----	-----	-----------------------	-----	-----	------------------------	-----	-----	------------------------	-----	-----	----------------------	-----	-----	-----------------------	-----	-----	----------------------	-----	-----	-----------------------	-----	-----	-----------------------	-----	-----	----------------------	-----	-----	-------------	-----	-------------

RESULT 8

ABU44499
ID ABU44499 standard; protein; 669 AA.

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #30026.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Streptococcus mutans.

XX
PN
WO200277183-A2.XX
PD 03-OCT-2002.XX
PF 21-MAR-2002: 2002WC-IIS009107XX
PB 21-MAR-2001: 2001US-0081524

PR 06-SEP-2001; 2001US-0094899
PR 25-OCT-2001; 2001US-0342923

PR 08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH; Zyskind L, Zamudio C, Malone C, Haselbeck R, Olsen K, Zyskind L, Zamudio C, Malone C, Haselbeck R, Olsen K, Xu HH; Xu HH;

WPI; 2003-029926/02.

WET, 2003-023320/
N-PSDB; ACA48369.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25: SEO ID NO 72423; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://www.int/pub/published/pct_sequences

Sequence 669 AA;

Query Match	82.9%	Score 2883	DB 6	Length 669
-------------	-------	------------	------	------------

82.9%; SCORE 2883; DB 8;
82.4%; Pred. No. 4.3e-234;

Sequence	Local Similarity	Indels	Mismatches	Gaps
1	82.4%	46	46	8
2	82.4%	46	46	8
3	82.4%	46	46	8
4	82.4%	46	46	8
5	82.4%	46	46	8
6	82.4%	46	46	8
7	82.4%	46	46	8
8	82.4%	46	46	8
9	82.4%	46	46	8
10	82.4%	46	46	8
11	82.4%	46	46	8
12	82.4%	46	46	8
13	82.4%	46	46	8
14	82.4%	46	46	8
15	82.4%	46	46	8
16	82.4%	46	46	8
17	82.4%	46	46	8
18	82.4%	46	46	8
19	82.4%	46	46	8
20	82.4%	46	46	8
21	82.4%	46	46	8
22	82.4%	46	46	8
23	82.4%	46	46	8
24	82.4%	46	46	8
25	82.4%	46	46	8
26	82.4%	46	46	8
27	82.4%	46	46	8
28	82.4%	46	46	8
29	82.4%	46	46	8
30	82.4%	46	46	8
31	82.4%	46	46	8
32	82.4%	46	46	8
33	82.4%	46	46	8
34	82.4%	46	46	8
35	82.4%	46	46	8
36	82.4%	46	46	8
37	82.4%	46	46	8
38	82.4%	46	46	8
39	82.4%	46	46	8
40	82.4%	46	46	8
41	82.4%	46	46	8
42	82.4%	46	46	8
43	82.4%	46	46	8
44	82.4%	46	46	8
45	82.4%	46	46	8
46	82.4%	46	46	8
47	82.4%	46	46	8
48	82.4%	46	46	8
49	82.4%	46	46	8
50	82.4%	46	46	8
51	82.4%	46	46	8
52	82.4%	46	46	8
53	82.4%	46	46	8
54	82.4%	46	46	8
55	82.4%	46	46	8
56	82.4%	46	46	8
57	82.4%	46	46	8
58	82.4%	46	46	8
59	82.4%	46	46	8
60	82.4%	46	46	8
61	82.4%	46	46	8
62	82.4%	46	46	8
63	82.4%	46	46	8
64	82.4%	46	46	8
65	82.4%	46	46	8
66	82.4%	46	46	8
67	82.4%	46	46	8
68	82.4%	46	46	8
69	82.4%	46	46	8
70	82.4%	46	46	8
71	82.4%	46	46	8
72	82.4%	46	46	8
73	82.4%	46	46	8
74	82.4%	46	46	8
75	82.4%	46	46	8
76	82.4%	46	46	8
77	82.4%	46	46	8
78	82.4%	46	46	8
79	82.4%	46	46	8
80	82.4%	46	46	8
81	82.4%	46	46	8
82	82.4%	46	46	8
83	82.4%	46	46	8
84	82.4%	46	46	8
85				

1 MSEKN-FYITTPITYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQ 59

1 MTEKQPFYITPTIYVPSGKLHIGSAYTTIACDVLARYKRMNHDVFYLTGLDEHGQIKQ 60

60 KAEFAGITPOAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAQDDIY 119

61 KSEEAGITPOAYVDCMAVGKELWLLDISYDKFIRTTDDYHEKWADVFEKLLAOGDIY 120

120 LGEYSGWYSVSDFFFFFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYOD 179

121 LGEYSGWYSVSDDEFFFTESOLEEIERDDEGKVIGGIAPSGHEVWVSEESYFLRLSNYAD 180

180 RI.VEEFKAHPEFTTPDGRI.NEMI.RNFEFPGLEDI.AVSR.TTFTWGVVPSPKHYVYVWID 239

181 PLYDEEHAHBPDEIOPDGRMNEITKNEIERGLEDI.AVSRTSEFTWGVKVPSPHWWYWD 240

240 3I.J.N.Y.8.TA1.G.Y.8.O.D.E.H.C.N.E.D.K.E.W.N.G.T.V.F.H.M.V.G.K.D.I.I.B.E.H.S.T.Y.W.P.T.I.I.M.M.I.D.V.K.I.P.D.R.I.T.A 299

241 21 INVA TAI CYCOP AHA NEP KE NC TVEH MW CK D II PEH ST YW B II I MM D I KM PER L A 300

CC 300 HGWFMKDKGMSKSGNVVPEMLVERGLDPLRYLLMRNLPGVSDGTFPTPEDYVGRIN 359
 CC 301 HGWFMKDKGMSKSGNVVPEMLVERGLDPLRYLLMRNLPGVSDGTFPTPEDYVGRIN 360
 CC 360 ELANDLGNLLNRTVSMINKYFGQIPAYVGVTEBDHVLAEVAKSIADPFTHEANDYP 419
 CC 361 ELANDLGNLLNRTVSMINKYFGQIPAYVGVTEBDHVLAEVAKSIADPFTHEANDYP 420
 CC 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMET 479
 CC 421 RALEVWNLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMET 480
 CC 480 SRAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPPIPRNLMBEIEIAYIKQWEGNKPA 539
 CC 481 SNAIMEQLGLAGQFLEGLLEAGLPTGVKVIAGKTPPIPRNLMBEIEIAYIKQWEGNKPA 539
 CC 540 V----EKWNPDEVELKLNKDSIKFEDPKVIRVAEVEKSVKSGDKLLQFRLDAGDG 595
 CC 540 IQAEEKEWNPPTVELKNEKKAIFDDFDKLEIRVAEVEKSVKSGDKLLKFLRLDAGDG 599
 CC 596 EDRQILSGIAKYYPNEQLVGGKQIVANLKPRKMMKKYVSGMILSAEH-DGKLTLLTV 654
 CC 600 EDRQILSGIAKYYPNEQLVGGKQIVANLKPRKMMKKYVSGMILSAEHADGKLTLLTV 658
 CC 655 DPAVNGSVIG 665
 CC 659 DSSVPNGSQIG 669

RESULT 9
 ABP30434
 ID ABP30434 standard; protein; 665 AA.
 XX ABP30434;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 10044.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB004789.
 XX
 XX 27-OCT-2000; 2000GB-00026333.
 XX
 XX 24-NOV-2000; 2000GB-00028727.
 XX
 XX 07-MAR-2001; 2001GB-000035640.
 XX
 XX (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 XX
 XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
 XX Tettelin H;
 XX
 XX WPI; 2002-352536/38.
 XX
 XX N-PSDB; ABN71065.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection or
 XX disease caused by Streptococcus bacteria, such as meningitis, and for
 XX detecting a compound that binds to the protein.
 XX
 XX Claim 1; Page 4133-4134; 4525pp; English.
 XX
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 665 AA;

Query Match 78.5%; Score 2732.5; DB 5; Length 665;
 Best Local Similarity 78.2%; Pred. No. 2.2e-221;
 Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
 3 EKRFYTTPIYPSGKHLHIGSAYTTTACDVLARYKRLMGYDFVLTGLDEHGGKIQQKAE 62
 4 KKSFYTTPIYPSGKHLHIGSAYTTTACDVLARYKRLMGYDFVLTGLDEHGGKIQQKAE 63
 63 EAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVVAOVFERLLAQDDIYLGE 122
 64 EAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVVAOVFERLLAQDDIYLGE 123
 123 YSCWYSVSDSEFFETESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRSKTKQDRILV 182
 124 YTCWYSVSDSEFFETESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRSKTKQDRILV 183
 183 EPPKAHPERTTPDRLNEMLRNFIPEGLDGLAVSRFTTGTWGPVSPNPKHVVYVWIDALL 242
 184 AYYAEHPERTTPDRLNEMLRNFIPEGLDGLAVSRFTTGTWGPVSPNPKHVVYVWIDALL 243
 243 NYATALGYA-QDEHGNFDKFWNGTVFHMVGKDLRPHSYIWPILMLMDVLPDLRIANG 301
 244 NYISALGYA-QDEHGNFDKFWNGTVFHMVGKDLRPHSYIWPILMLMDVLPDLRIANG 302
 302 WFWMDGKMSKSGNVVPEMLVERGLDPLRYLLMRNLPGVSDGTFPTPEDYVGRINVEL 361
 303 WFWMDGKMSKSGNVVPEMLVERGLDPLRYLLMRNLPGVSDGTFPTPEDYVGRINVEL 362
 362 ANDLGNLLNRTVSMINKYFGQIPAYVGVTEBDHVLAEVAKSIADPFTHEANDYPRA 421
 363 ANDLGNLLNRTVSMINKYFGQIPAYVGVTEBDHVLAEVAKSIADPFTHEANDYPRA 421
 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMETSR 481
 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDQLASVNSHWQASIRVVAHLIPEPFMMETSD 481
 482 AVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPPIPRNLMBEIEIAYIKQWEGNKPAVE 541
 482 AIMEQLGLAGTFLDKLTPADLPBGVTVVAKGSPIPRNLMBEIEIAYIKQWEGNKPAVE 541
 542 KWNPEVELKLNKDEIKFEDPKVIRVAEVEKSVKSGDKLLQFRLDAGDGQRIL 601
 542 KWNPEVELTSSKGQIKFEDPKVIRVAEVEKSVKSGDKLLQFRLDAGDGQRIL 601
 602 SGIAKYYPNEQLVGGKQIVANLKPRKMMKKYVSGMILSAEHDKLTLLTVDPVANG 661
 602 SGIAKYYPNEQLVGGKQIVANLKPRKMMKKYVSGMILSAEHDKLTLLTVDPVANG 661
 662 SVIG 665
 662 SIIG 665

ID ABP26525 standard; protein; 682 AA.
 XX AC ABP26525;
 XX DT 02-JUL-2002 (first entry)
 XX DE Streptococcus polypeptide SEQ ID NO 2226.
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX OS Streptococcus agalactiae.
 XX WO200234771-A2.
 XX PD 02-MAY-2002.
 XX PF 29-OCT-2001; 2001WO-GB004789.
 XX PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Masignani V, Margarit Y RosI, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN67156.
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX Claim 1; Page 3374; 4525pp; English.
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX Sequence 682 AA;
 Query Match 78.5%; Score 2732.5; DB 5; Length 682;
 Best Local Similarity 78.2%; Pred. No. 2.3e-221;
 Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
 QY 3 EKNEYITPIYPSGKLHIGSAYTTIACDVLARKLGYDVYVLTGLDHEGKIOKAE 62
 DB 21 KKSFYITPIYPSGKLHIGSAYTTIACDVLARKLGYDVYVLTGLDHEGKIOKAE 80
 QY 63 EAGITPOAYVDGMAVGKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAQDDIYLGE 122
 DB 81 EAGITPOEYVDGMAESVKLTWELLDDISYDKFIRTTDYHEEAVAKIFEQLAQDDIYLGE 140
 QY 123 YSGWYSVDEBFFTESQLAEVYRDENGNMIGVAPSGHEVKEVSESYFFRMSKYADELK 182

DB 141 YTGWYSVDEBFFTESQLAEVYRDENGNMIGVAPSGHEVKEVSESYFFRMSKYADELK 200
 QY 183 EFKKAHPEFITPDGRLNEMLRNFTEPGLDLAVGRTTFTWGPVPSNPKHVVYVWIDALL 242
 DB 201 AYAHEPEFIQPDGRNEMLRNFTEPGLDLAVGRTTFTWGPVPSNPKHVVYVWIDALM 260
 QY 243 NYATALGYA-QDEHGNFDFKFWNGTVFHMVGDIIRFHSIYWPILLMLMDVKLPDLRIAHG 301
 DB 261 NYISALGYCWSDDLSSQYHKFPADI-EMIGKDIIRFHSIYWPIMLMALDLPKRLVAHG 319
 QY 302 NFWMKDGKWSKGNVYVPEMLVERYGLDPLRYILMRNLVPGSDGTFTPEDYVGRINVEL 361
 DB 320 NFWMODGKWSKGNVYVPEMLVERFGLDPLRYILMRSLVPGSDGTFTPEDYVGRINVEL 379
 QY 362 ANDLGNLNRITVSMINKYFDQIQIPAYVEGTVTFDHVLAEVAEKSIADPHTHMEAVDYFPA 421
 DB 380 ANDLGNLNRITAMVNYKFDGEVPRFAV-ATDFDADLASVATDSTENYHKQMEAVDFPFA 438
 QY 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHWOASIRVVVAHLIEPFWMTSR 481
 DB 439 LEAVWNLISRTNKYIDETAPWLVAKDETDRDKLAAMSHLVASURVVAHLIQPFWMTSD 498
 QY 482 AVLTOIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAVE 541
 DB 499 AIMEQLGLGATFDEKLTFAADLPEGVRVVAKSGPIFPFLDMEDEITYIKEQMNAGKAPVE 558
 QY 542 KEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSKLLQPLDAGDGEDROIL 601
 DB 559 KEWPEVEVELTSSKGQIKFEDFAYEIRVAEVEKVEGSKLLRFLDAGDEHQRIL 618
 QY 602 SGIAKYVNEOELVGKKVQIVANLKPROMKKYVSOQGMILSAEHDGKLTLLTVDPAVNG 661
 DB 619 SGIAKYVNEOELVGKKLQIVANLKPROMKKYVSOQGMILSAEHDGKLTLLTVDPAVNG 678
 QY 662 SVIG 665
 DB 679 SIIG 682
 RESULT 11
 ID ABB54105
 XX ABB34105 standard; protein; 662 AA.
 AC ABB34105;
 XX DT 29-AUG-2003 (revised)
 DT 16-MAY-2002 (first entry)
 DE Lactococcus lactis protein metS.
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 OS Lactococcus lactis; IL1403.
 XX FR2807446-A1.
 XX PD 12-OCT-2001.
 PF 11-APR-2000; 2000FR-00004630.
 XX 11-APR-2000; 2000FR-00004630.
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 PI WPI; 2002-043418/06.
 XX New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species.
 XX Claim 6; SEQ ID NO 807; 2504pp; French.

CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
CC acid sequence is useful in the detection and/or amplification of nucleic
CC acid sequence, particularly to identify Lactococcus lactis or related
CC species. The proteins of the invention are useful for the biosynthesis or
CC biodegradation of a composition of interest. The invention helps research
CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)

XX SQ Sequence 662 AA;

Query Match 68.8%; Score 2392.5; DB 5; Length 662;
Best Local Similarity 68.8%; Pred. No. 1.1e-192;
Matches 458; Conservative 88; Mismatches 113; Indels 9; Gaps 7;

QY 1 MSF-KNPIYITPTIYPSGKLIHGSAYTTIACDVLARYKELMGDYVYLYTLGLDGHGQKIQ 59
DB 1 MTENKTFYITPTIYPSGKLIHGSAYTTIACDVLARYKELMGDYVYLYTLGLDGHGQKIQ 60
QY 60 KAEAGITTPQAYVDGMAVGKELMQLLDISYKFIKRTTDDYHEKVAQVFERLLAQDDIY 119
DB 61 KAEELGWTPEKYLDPVADVOELMKKLDISYKFIKRTTDDYHEEAVAKAFELQLDQDIY 120
QY 120 LGSYGWSVSDDEFTTESQLAEVFRDEAGNVTGGIAPSGCHEVWVSESYFLRSKYQD 179
DB 121 LGKYAGWSVSDDEFTTESQLAEVFRDEAGNVTGGIAPSGCHEVWVSESYFLRSKYQD 180
QY 180 RLVFEKAPHEFTTPDGRNLEMLRNPIEGLEDLAVSRFTFTGVPVPSNPKHVVYVWD 239
DB 181 WLLQYVDHDFDQPEVRKENVNPNFKGLEBDLALTRTSFTGVPVPSNPKHVVYVWD 240
QY 240 ALLNVALYALGYAQDEHGFNFKWNGVFMVGVKDLRHSIYWPILLMMLDKLPLDLIA 299
DB 241 ALLNVALYALGYNSDNDNSNFKYWP--INMGKEIVRFTIYWPIMLHALGPAKKIFA 298
QY 300 HGFVMDKQWSKSGNVVYPEMLVRYGLDPLRYLYMLRNLPVSGDGTTPEDYVGRNY 359
DB 299 HGWLLMKDGRWSKSGNVVYPEMLVRYGLDPLRYLYMLRNLPVSGDGTTPEDYVGRNY 358
QY 360 ELANDLGNLLNRVSMINKYFDGCIYAVYGVG--TEPDHVLAEVAKSIADFTTHMEAVDY 418
DB 359 DLANDLGNLLNRVSMINKYNDGKIEA--TGVSFEDASLEEVETISHFKAMDKEF 416
QY 419 PRALEAVWTLISSTNKYIDETAPWLDKDBALRDQLASVNSHWQASIRVVAHLIPEFME 478
DB 417 NVALADVWTLISSTNKYIDETAPWLDKDBALRDQLASVNSHWQASIRVVAHLIPEFME 476
QY 479 TSRAVLTOIGLEVS--SLENLSLADPPADVTVVAKTPIPLRNMEETAYIKMEGK 537
DB 477 TSKKITEQLGMDRSFLENLSLSP--YSFTHPVVAKGQPIPPRLDVEEVAYIKLQVAGG- 534
QY 538 PAYEKEWNPDEVELKLNKDBIKPEDFKVEIRVAEVKEYSKVGEKDLQPLDAGDGED 597
DB 535 VLPEKEWNPDEVELNLTLPQIKFDDFEKIELKVALEVEPEVSGDKLRLFRKLDAGDSEP 594
QY 598 RQLSLGIKYPNEQLVGVKQVIVANLPRKMMKYYVSGMILSAEHDKLTLTVDPA 657
DB 595 RQLSLGIKYPNEQLVGVKQVIVANLPRKMMKYYVSGMILSAEHDKLTLTVDPA 654
QY 658 VPNGSVIG 665
DB 655 VPAGSLIG 662

RESULT 12
ABU29885
ID ABU29885 standard; protein; 669 AA.
XX AC ABU29885;
AC ABU29885;

XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by prokaryotic essential gene #15412.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Enterococcus faecium.
XX FN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362659P.
XX FA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WIPI: 2003-029926/02.
XX DR N-PSDB; ACA33755.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX FS Claim 25; SEQ ID NO 57809; 1765pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 669 AA;

Query Match 66.0%; Score 2295; DB 6; Length 669;
Best Local Similarity 65.2%; Pred. No. 1.9e-184;
Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;

DB 488 IFNQLGSEWNLEGLHGFPPSGTKVAVKGTPIFPRLDMEEEVAFIQEKMSSEGTQNE 547
QY 538 FAVEKEMPDPDEVELKLNKD-EIKFDFDKVEIRVAEYKESKVEGSKLLQFRILDAGDE 596
DB 548 DTV--KMDPEETELVSTKEKQIKFDVFEKVELKVAEIVNCQKVEGADKLLQFRILDAGDSQ 605
QY 597 DRQILSGIAKYPNPQBELVGKQVIVANLKRKMKKVVSGMILSAB-HGKLTLLTVD 655
DB 606 DRQILSGIAEFYDPSELIGKQWIVANLKRK-MRQGISQGMILSABAPGSLQVIEAP 664
QY 656 PAVPNGSVI 664
DB 665 KMPNGSEI 673

RESULT 14

AAU35070
ID AAU35070 standard; protein; 677 AA.

XX AC AAU35070;

XX DT 13-FEB-2002 (first entry)

XX DE Enterococcus faecalis cellular proliferation protein #357.

XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;

XX KW antibacterial; drug design.

XX OS Enterococcus faecalis.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009180.

XX PR 21-MAR-2000; 2000US-0191078P.

XX PR 23-MAY-2000; 2000US-0206848P.

XX PR 26-MAY-2000; 2000US-0207272P.

XX PR 23-OCT-2000; 2000US-0242578P.

XX PR 27-NOV-2000; 2000US-0253625P.

XX PR 22-DEC-2000; 2000US-0257931P.

XX PR 16-FEB-2001; 2001US-0269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX DR WPI; 2001-611495/70.

XX DR N-PSDB; AAS52929.

XX PT New polynucleotides for the identification and development of

XX PT antibiotics, comprise sequences of antisense nucleic acids.

XX PS Example 3; SEQ ID NO 10663; 511pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes,
XX CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
XX CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
XX CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 677 AA;

Query Match 65.8%; Score 2288; DB 4; Length 677;
Best Local Similarity 63.9%; Pred. No. 7, 7e-184;
Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYTTTPIYPSGKLHIGSAYTTIACDLARYKLMGYDYFYLTLGDEHQKIQQ 59
DB 9 MSQKETFYTTTPIYPSGKLHIGSAYTTIACDLARYKLMGYDYFYLTLGDEHQKIEK 68
QY 60 KAEERAGITPOAYVDGMAVGKELWQLDISYDKFTRTTDDYHEKVVAOVFERLLAQDDIY 119
DB 69 KAAELNVTPEYVDKMAADVQKLWLTLDISYDKFTRTTDDYHMAAVQQIFDELVEQSDIY 128
QY 120 LGSEYSGWYSVDEEFTTESQLAEVFRDAGNVTGGIAPSGHEVWVSSEYFLRLSKYQD 179
DB 129 LGSEYSGWYSVDEEFTTESQLAEVFRDAGNVTGGIAPSGHEVWVSSEYFLRLSKYAD 188
QY 180 RLVEPFKAHPEITPDGRLEMLRNFIIEGLEDLAVSRITFTMGVPVSPNPKHVVYVMD 239
DB 189 RLQVYEEHPEFIQESRKNEMLNFIKPLEDLAVSRITFTMGVPVSPNPKHVVYVMD 248
QY 240 ALINATATGVAQDEHGNFDFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
DB 249 ALSNYITALGYGSEDDSLFKYWPANV-QMVGKEIVRPHIYWPIMLMALDPLPKVFG 307
QY 300 HGWFWKMGKSKSGNVVYPEMLYERGLDPLRYLMLNLPVSGDGTTPEDYVGRINY 359
DB 308 HGWLLMKDGKSKSGNVVYPEMLYERGLDPLRYLMLNLPVSGDGTTPEDYVGRINY 367
QY 360 ELANDLGNLNLRTVSMINKYFDGQIPAYVEGTEFDHVAEVAEKSIADFFHMEAVDYP 419
DB 368 DLANDLGNLNLRTVSMINKYCDGKVPAYASKYTPFDSLSLTTAANVIGYHEAMEKMEFN 427
QY 420 RALEAVWTLISRTNKYIDTAPWLDKDEALDQDLASVMSHWQASIRVVAHLIESFMET 479
DB 428 TATAEITWLVSRANKYIDTAPWLVAKKEKEKELESVMHIALSLRIVAILLOQVMTET 487
QY 480 SRAVLTCIGLE-EVSSLENLSLADPADVTYVAKGTPFPPLNMEELIAYIKOEQGNKP 538
DB 488 PGKIFQLGLDPEPMNENIHFGFEPDVTYVTSKGTPIFPRLEITEVTYIQQKMSQSES 547
QY 539 AVEKE--WNPDEVELKLNKD-EIKFDFDKVIRVAEYKESKVEGSKLLQFRILDAGD 595
DB 548 ATEEDIKNPEETTLVSTKEKQIKYDDDFDKVELKVAEVIDCKKVGADKLLQFRILDAGD 607
QY 596 EDRQILSGIAKYPNPQBELVGKQVIVANLKRKMKKVVSGMILSABH-DGKLTLLTV 654
DB 608 NHRQILSGIAEFYDPSELIGKQWIVANLKRK-MRQGISQGMILSABSPGSLQVIEA 666
QY 655 DPAVNGSVI 664
DB 667 PREMPNGAGI 676

RESULT 15

ABU14551
ID ABU14551 standard; protein; 677 AA.

XX AC ABU14551;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by prokaryotic essential gene #78.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Enterococcus faecalis.

XX PN WO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-C0815242.

06-SEP-2001; 2001US-C0948993.

25-OCT-2001; 2001US-C342923P.

08-FEB-2002; 2002US-C00072851.

06-MAR-2002; 2002US-C0326295P.

(ELITR-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

DR N-PSDB; ACA18421.

DR

XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

PT

XX

PS Claim 25; SEQ ID NO 42475; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of the 5213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct_sequences](http://wipo.int/pub/published/pct_sequences)

XX

Sequence 677 AA;

XX

	Query Match	65.8%	Score 2288;	DB 6;	Length 677;	
	Best Local Similarity	63.9%;	Pred. No. 7.7e-184;			
	Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7					
Qy	1	MSEK-NFYITTPYYPSGKLHIGSAYTTIACDVLARYKRLMGVDYVFLTGLDEHGKIQQ	59			
		:::::	:::::			
		:::::	:::::			
Dd	9	MSQKETFYITTPYYPSGKLHIGNSYTTIACDAMARYKRLMGDFVFLTVLGDEHGQKIEK	68			
		:::::	:::::			
Qy	60	KAEAGIIPQAIVDGMAVGVKELWQLLDISYDKFIRTTDDYHEKKVAQVERLLAQQDIY	119			
		:::::	:::::			
Dd	69	KAAELNVTPEKYVDQWAADVOKLWKTLDISYDKFIRTTDDYHMAAQVFDRLVEGGDIY	128			
		:::::	:::::			
Qy	120	LGEYSGWTYSVDSBEFFTSQIAEVPFBAGNVNTGGIAPSPGHEVETWBSGSYFURLSKYQD	179			
		:::::	:::::			

129	LEYEGWTSVSDSEBFF	TETQIAEYVRDDEGNVIGGKAPSGHEVELVKEBSYFFRMSKYAD	188
180	RLVPEFKAHPEFI	TPDQGLNEMLNFIPEGLDLAVSRFTTGWGPVPSNPKHVWVWID	239
189	RLVQVYSEHPFI	QPESKNEMLNFIKPGLEDLAVSRFTSGWIPLNDKPHVWVWID	248
240	ALLNVATALGVAQDE	HGNFDKFWNGTVFHMVGKDILRFHSIYWPILLMLDVKLPDLRIA	299
249	ALSNITVALGVSEDS	SLFOKYWPNV-QWVGKEIVRFHTIYWPILMLDULPLPKKVFVG	307
300	HGWFWKDGKMSKSGN	VVYPEMILVERYGIDPLRYILMNLPLVGSOGTFTTPEDYVGRINY	359
308	HGWLLMKDGKMSKSGN	VVYPEMILVERYGIDALRYILLRAIPFGSDGVFTPEDFVSRNLN	367
360	ELANDLGNLLART	YSMINKYPDGOIPAVVEGVTEFDHVLAEVKAESIDAPHTHMEAVDP	419
368	DIANDLGNLLART	IAMINKYCDGKVPAYAKVTPFDSELSSTTAAVIGKXHEAMKEWEN	427
420	RALEAVWTLIS	TNKYIDETAPWVLKDDEALRDOLASVMSHWQASIRVVAHLIEPPMMET	479
428	TAIAEITWLVSRANKY	IDETAPWVLAKKEEKRENELESWTHLAESIRVAILLOPVMTET	487
480	SRAVLTOIGLE-EV	SSLENLSLADFPADTVVAKGTPIFPRLNWEEBIAVIKOMEGNKP	538
488	PKGIEQIGLOPETH	NMENIHFGFPDVTIVTSKGTPIFPRLEIETEVYIQKNSQSES	547
539	AVEKE--WNPEVEL	KUNKO-EIKFEPDVKVETIRVAEVEKVSKEGSKDLQPLRDAGOG	595
548	ATEEDIKWNPETH	VLSTKEXQIYDDPDKVELKVAEVIDCKKVKGADKLLQPLRDAGDE	607
596	EDRQILSGIAKYP	NEQELVGKKVQIVANLKPRKMKKVVYSQGMILSAEH-DGKLLTLTV	654
608	NHRQILSGIAEY	PDPAALIGKKVVIIVANLKPRK-MEQOISQGMILSASPEKQIIVERA	666
655	DPAPVNGSVI	664	
667	PKEMPNGAGI	676	

Search completed: December 17, 2004, 21:56:06
Job time : 164 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:51:16 ; Search time 40 Seconds
(without alignments)
1102.537 Million cell updates/sec

Title: US-10-606-607-2

Perfect score: 3479

Sequence: 1 MSEKNFYITPTIYPSGKLH.....DGKLTLLVDPAPNGSVIG 665

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/aaa/5A-COMB.pep:.*
2: /cgn2_6/prodata/1/aaa/5B-COMB.pep:.*
3: /cgn2_6/prodata/1/aaa/6A-COMB.pep:.*
4: /cgn2_6/prodata/1/aaa/6B-COMB.pep:.*
5: /cgn2_6/prodata/1/aaa/PTUS-COMB.pep:.*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3479	100.0	665	3	US-08-844-059-2
2	3479	100.0	665	3	US-09-431-202-2
3	3429	98.6	679	4	US-09-583-110-2751
4	2291.5	65.9	674	4	US-09-107-532A-6201
5	1922.5	55.3	660	3	US-09-134-001C-3350
6	1767.5	50.8	618	4	US-09-710-279-860
7	1492.5	42.9	493	4	US-09-134-000C-4034
8	1137.5	32.7	648	1	US-08-451-715A-4
9	1033	29.7	521	1	US-08-584-226-2
10	755	21.7	683	4	US-09-543-681A-5403
11	726	20.9	716	4	US-09-489-039A-8243
12	701.5	20.2	402	4	US-09-248-796A-16768
13	661	19.0	720	4	US-09-328-352-4765
14	633.5	18.2	690	4	US-09-540-236-3507
15	625	18.0	680	4	US-09-252-991A-17566
16	619	17.8	382	4	US-09-270-767-62086
17	614	17.6	127	3	US-08-858-207A-490
18	574	16.5	115	3	US-08-844-059-4
19	574	16.5	115	3	US-09-431-202-4
20	493	14.2	554	4	US-09-198-452A-140
21	470.5	13.5	659	3	US-09-392-772-10
22	431.5	12.4	176	4	US-09-710-278-2332
23	344	9.0	152	3	US-09-392-772-12
24	310	8.9	888	3	US-09-134-001C-3032
25	308.5	8.9	876	1	US-08-785-429-2
26	308.5	8.9	876	3	US-08-996-621-2
27	300	8.6	883	4	US-09-583-110-2900

28 299 8.6 883 2 US-08-953-492-2 Sequence 2, Appli
29 296 8.5 902 4 US-09-107-532A-3837 Sequence 3837, Ap
30 286 8.2 894 4 US-09-134-000C-5731 Sequence 5731, Ap
31 265.5 7.6 80 4 US-09-134-000C-3795 Sequence 3795, Ap
32 256.5 7.4 804 1 US-08-785-428-2 Sequence 2, Appli
33 256.5 7.4 804 2 US-08-996-797-2 Sequence 2, Appli
34 240 6.9 128 4 US-09-270-767-62086 Sequence 62086, A
35 236 6.8 872 1 US-08-451-715A-8 Sequence 8, Appli
36 233.5 6.7 835 3 US-09-134-001C-5105 Sequence 5105, Ap
37 236.5 6.5 1045 1 US-08-452-083-2 Sequence 2, Appli
38 220.5 6.3 899 3 US-09-413-814-5 Sequence 5, Appli
39 216 6.2 318 4 US-09-248-796A-18266 Sequence 18266, A
40 213 6.1 824 4 US-09-134-000C-4908 Sequence 4908, Ap
41 212.5 6.1 833 2 US-08-844-086-2 Sequence 2, Appli
42 212.5 6.1 833 3 US-09-018-211-2 Sequence 2, Appli
43 212 6.1 804 3 US-08-855-910-4 Sequence 4, Appli
44 210.5 6.1 833 4 US-09-583-110-4098 Sequence 4098, Ap
45 205.5 5.9 828 4 US-09-107-532A-4191 Sequence 4191, Ap

ALIGNMENTS

RESULT 1
US-08-844-059-2
; Sequence 2, Application US/08844059
; Patent No. 6001601
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NO. 6001601el Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,059
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607999.1
; FILING DATE: 16-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-844-059-2

Query Match 100.0%; Score 3479; DB 3; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-303;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSEKNFYITPTIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVYLTGLDHGQKIQK 60
|||||

Db 1 MSEKMFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVYLTGLDEHGKIQOK 60
QY 61 ABEAGITPOAYVDGMAGVVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPOAYVDGMAGVVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSGWYSVSDDEEFTESQLAVERDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR 180
Db 121 GEYSGWYSVSDDEEFTESQLAVERDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNRLNFIIEPGLDLAVSRITFTWGVVPSNPKHVVVWIDA 240
Db 181 LVEFFKAHPEFITPDGRNLNRLNFIIEPGLDLAVSRITFTWGVVPSNPKHVVVWIDA 240
QY 241 LLYATATGAYQADEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
Db 241 LLYATATGAYQADEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPEDYVGRINYE 360
Db 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPEDYVGRINYE 360
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFFTHMEAVDPR 420
Db 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFFTHMEAVDPR 420
QY 421 ALEAVWTLISRTNKYIDETA PWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMETS 480
Db 421 ALEAVWTLISRTNKYIDETA PWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEIAIYKEQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEIAIYKEQMEGNKPAV 540
QY 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQRLDAGGEDRQI 600
Db 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQRLDAGGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 601 LSGIAKYYPNEQELVGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

RESULT 2

US-09-431-202-2
; Sequence 2, Application US/09431202
; Patent No. 6294175
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6294175e1 Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,202
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,059

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: F31456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-202-2

Query Match 100.0%; Score 3479; DB 3; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-303;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSEKMFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVYLTGLDEHGKIQOK 60
Db 1 MSEKMFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVYLTGLDEHGKIQOK 60
QY 61 ABEAGITPOAYVDGMAGVVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPOAYVDGMAGVVKELWQLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYL 120
QY 121 GEYSGWYSVSDDEEFTESQLAVERDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR 180
Db 121 GEYSGWYSVSDDEEFTESQLAVERDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPDGRNLNRLNFIIEPGLDLAVSRITFTWGVVPSNPKHVVVWIDA 240
Db 181 LVEFFKAHPEFITPDGRNLNRLNFIIEPGLDLAVSRITFTWGVVPSNPKHVVVWIDA 240
QY 241 LLYATATGAYQADEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
Db 241 LLYATATGAYQADEHGNFKWNGTVFHMVGKDIILRFHSIYWPILLMMLDVKLPDLRIAH 300
QY 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPEDYVGRINYE 360
Db 301 GWFVMDGKMSKSGNVVYPEMLVERYGLDPLRYLLMRNLFPVSGDGTFTPEDYVGRINYE 360
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFFTHMEAVDPR 420
Db 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADFFTHMEAVDPR 420
QY 421 ALEAVWTLISRTNKYIDETA PWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMETS 480
Db 421 ALEAVWTLISRTNKYIDETA PWLDKDEALRDQLASVMSHWQASIRVVAHLIEPMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEIAIYKEQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIFPLNMEEBEIAIYKEQMEGNKPAV 540
QY 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQRLDAGGEDRQI 600
Db 541 EKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQRLDAGGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 601 LSGIAKYYPNEQELVGKQVIVANLKPROMKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

RESULT 3

US-09-583-110-2751

; Sequence 2751, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 05/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2751
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2751

Query Match 98.6%; Score 3429; DB 4; Length 679;
Best Local Similarity 98.8%; Pred. No. 5,1e-299;
Matches 657; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQK 60
Db 15 MSEKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQK 74

QY 61 ABEAGITPOAVDGMVGVKELQQLDISYDKFIRITDDYHEKVAQVFERLLAQDDIYL 120
Db 75 ABEAGITPOAVDGMVGVKELQQLDISYDKFIRITDDYHEKVAQVFERLLAQDDIYL 134

QY 121 GEYSGWTSVSDDEEFFTESQLAEVRDEAGNVTGGIAPSGHEVWVSESYFLLSLKYQDR 180
Db 135 GEYSGWTSVSDDEEFFTESQLAEVRDEAGNVTGGIAPSGHEVWVSESYFLLSLKYQDR 194

QY 181 LVEFFKAHPFIFPDGRLNEMLANFIEPGLDIAVSRITTTGVPVPSNPKHVVYVWIDA 240
Db 195 LVEFFKAHPFIFPDGRLNEMLANFIEPGLDIAVSRITTTGVPVPSNPKHVVYVWIDA 254

QY 241 LLNVATAGVQAQDHGDFWNGTVFHMVGKIDLRPHSIYWPILLMLLDVKLPDRLLAH 300
Db 255 LLNVATAGVQAQDHGDFWNGTVFHMVGKIDLRPHSIYWPILLMLLDVKLPDRLLAH 314

QY 301 GWFYKDGKSKSGNVVYVPEMLVERYGDLPLRYLLMRNLPGSDGTFTPEDYVGRINYE 360
Db 315 GWFYKDGKSKSGNVVYVPEMLVERYGDLPLRYLLMRNLPGSDGTFTPEDYVGRINYE 374

QY 361 LANDGLNLRNRTSMINKYFDGQIPAYVEGVTBFDHVAEVAEKSIADEFTHMEAVDYPR 420
Db 375 LANDFGNLRNRTSMINKYFDGQIPAYVEGVTBFDHVAEVAEKSIADEFTHMEAVDYPR 434

QY 421 ALEAVMTLISRTNKYIDETAPVLDKDEALRDQALASVMSHWQASIRVAHLIEPFMMETS 480
Db 435 ALEAVMTLISRTNKYIDETAPVLDKDEALRDQALASVMSHWQASIRVAHLIEPFMMETS 494

QY 481 RAVLTQGLBEVSLNLSADPADVTVAKTPIFPLNMBEEIATYKEQVEGNKPAV 540
Db 495 RAVLTQGLBEVSLNLSADPADVTVAKTPIFPLNMBEEIATYKEQVEGNKPAV 554

QY 541 EKEWNPDEVELKLNKDRIKPEDFKVEIRVAEVEKSVKVGSKLLQFLDADGEDRQI 600
Db 555 EKEWNPDEVELKLNKDRIKPEDFKVEIRVAEVEKSVKVGSKLLQFLDADGEDRQI 614

QY 601 LSGIAKYPNQEVLGKKVQIVANLKPROMKKYVSGMILSAEHGKTLTLLTVDPAPVN 660
Db 615 LSGIAKYPNQEVLGKKVQIVANLKPROMKKYVSGMILSAEHGKTLTLLTVDPAPVN 674

QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 4

US-09-107-532A-6201
; Sequence 6201, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arintello, Pamela Denke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6201:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...674
; SEQUENCE DESCRIPTION: SEQ ID NO: 6201:
US-09-107-532A-6201

Query Match 65.9%; Score 2291.5; DB 4; Length 674;
Best Local Similarity 65.0%; Pred. No. 7,5e-197;
Matches 435; Conservative 94; Mismatches 129; Indels 11; Gaps 8;

QY 3 EKNFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQKAE 62
Db 9 KETFYITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVYLTGLDEHGQKIQQKAE 68

QY 63 EAGITPOAVDGMVGVKELQQLDISYDKFIRITDDYHEKVAQVFERLLAQDDIYLGE 122
Db 69 ELGVKPEYVDKMAADVKKLWKTLDISYDKFIRITDDYHKAQVQKIFRLEQGDIIYLGE 128

QY 123 YSGWTSVSDDEEFFTESQLAEVRDEAGNVTGGIAPSGHEVWVSESYFLLSLKYQDRLV 182
Db 129 YEGWTSVSDDEEFFTESQLAEVRDEAGNVTGGIAPSGHEVWVSESYFLLSLKYQDRLV 188

QY 183 BEFKAHPSFIPTDGRNEMLANFIEPGLDIAVSRITTTGVPVPSNPKHVVYVWIDALL 242
Db 189 EYEEHPFIQPSRKNNINFIKPLGLEDIAVSRITTTGVPVPSNPKHVVYVWIDALLS 248

243 NYATGALGAQDEHGNFDFWNGTVPFHMVGKDIIRPHSYWPIILLMMLDVKLPDRLLAHGW 302
249 NYITAGLGSDESSEFEKIWPADV-QMVGKELVRHTIYWIIMLMALDLPFKKVFHGW 307
303 FYMKDGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLVPGSDGTFPTEDVYGRINVELA 362
308 LLMKDGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLVPGSDGTFPTEDVYGRINVELA 367
363 NDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHEAVDYPAL 422
368 NDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHEAVDYPAL 427
423 EAVWTLISRTNKYIDETAPWLDKDEALRDQLASVSHWQASIRVVAHLIIPPMMETSA 482
428 AETWLVISRTNKYIDETAPWLDKDEALRDQLASVSHWQASIRVVAHLIIPPMMETSA 487
483 VLTQIGLE-EVSSLENLSIADFPADVTVVAKGTPFPRLNMEETIAYIKQM-EG--NK 537
488 IFNLGLDSETNLEGLFGEFPGPKVAKGTPFPRLNMEETIAYIKQM-EG--NK 547
538 PAVEKEWNPDEVELKLNK-D-EIKFEDFDKVEIRVAEVEKSVKVEGSDKLLQFRLDAGDGE 596
548 DTV--KWDPEETELVSTKQIKFDFVEKVELKVAEINQCKVEGADKLLQFRLDAGDSQ 605
597 DRQILSGIAKYPNEQELVCKKQVIVANLKPRKMMKKYVSGMILSAE-HDCKLTLTVD 655
606 DRQILSGIAEFYDPSELGKVKVIVANLKPRK-MRGQISOGMILSAEAPDGSLOVIEAP 664
656 PAVPNGSVI 664
665 KSPNGSEI 673

RESULT 5

US-09-134-001C-3350
; Sequence 3350, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3350
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3350

Query Match 55.3%; Score 1922.5; DB 3; Length 660;
Best Local Similarity 55.9%; Pred. No. 1e-163;
Matches 375; Conservative 113; Mismatches 160; Indels 23; Gaps 11;
1 MSEKNFYITPIYPSGKLHIGSAITACDLARYRLMGVDFVLTGLDGHGKIQK 60
5 MAKDTFYITPIYPSGKLHIGSAITACDLARYRLMGVDFVLTGLDGHGKIQK 64
61 ABEAGITPQAYVDMGAVGVKELQLLDISYDKIRITDDVHEKVAQVFERLLAODDIYL 120
65 AQAAGKTELEYLDMSIGLKNLMSKLEISNDFFIRITEERKKQVKEVFERLLKQDIYL 124
121 GEYSGMYSVDEEFPFESQLAE-VFRDEAGNVTGGIAP-SGHEVWVSSESYFLRLSKYQ 178
125 GEYSGMYSVDEEFPFESQLAE-VFRDEAGNVTGGIAP-SGHEVWVSSESYFLRLSKYQ 182
179 DLVEFFKAHPEFTTPDGRNLMRLNIEQLDGLAVSRITFTTGWGVPSNPKHVVYVVI 238

183 DRLEFYDENPFIQPPSRKNEMINNFIKPGLEDIAVSRSTFDWGVVRPSNPKHVVYVVI 242
239 DALLNYATGALGAQDEHGNFDFWNGTVPFHMVGKDIIRPHSYWPIILLMMLDVKLPDRLL 298
243 DALVNYISSILGYLSDDETLPFKYWPADI-HLMAKEIVRPHSIIPWILLMALDLPFKKVF 301
299 AHGWFVMKDGKMSKSGNVVYPEMLVERGLDPLRYLLMRNLVPGSDGTFPTEDVYGRIN 358
302 AHGWLIMKDGKMSKSGNVVDPNVLIDRYGLDPLRYLLMRNLVPGSDGTFPTEDVYGRIN 361
359 YELANDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHEAVDYPAL 418
362 YELANDLGNLNTVSMINKYFDGQIPAYVEGTEPDHVLAEVAEKSIADFTHEAVDYPAL 421
419 PRALAEVWTLISRTNKYIDETAPWLDKDEALRDQLASVSHWQASIRVVAHLIIPPMMETSA 478
422 SVALSTWKFISRTNKYIDETAPWLDKDEALRDQLASVSHWQASIRVVAHLIIPPMMETSA 481
479 TSRAVLTOIG-----LBEVSSLENLSIADFPADVTVVAKGTPFPRLNMEETIAYIKQM 533
482 APREIFKQNLNPNPDHLQDLSLQYGM--LSEAITVTEKTPFPRLDTEAEIAYIKESM 539
534 EGNKPAVEKEWNPDEVELKLNK-D-EIKFEDFDKVEIRVAEVEKSVKVEGSDKLLQFRLDAG 593
540 QPK-SIKQSDPE-----GKEQIDIKDFKVEIRVAEVEKSVKVEGSDKLLQFRLDAG 590
594 DGEDRQILSGIAKYPNEQELVCKKQVIVANLKPRKMMKKYVSGMILSAE-HDCKLTLTVD 653
591 DNEQRQIVSGIAKYPNEQELVCKKQVIVANLKPRKMMKKYVSGMILSAE-HDCKLTLTVD 648
654 VDAVPNGSVI 664
649 LPSAIPNGAVI 659

RESULT 6

US-09-710-279-860
; Sequence 860, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 860
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-860

Query Match 50.8%; Score 1767.5; DB 4; Length 618;
Best Local Similarity 54.8%; Pred. No. 7.7e-150;
Matches 347; Conservative 108; Mismatches 155; Indels 23; Gaps 11;
39 LMGYDFVLTGLDGHGKIQKAEAGITPQAYVDMGAVGVKELWQLLDISYDKIRITDD 98
1 MOGYDFVLTGLDGHGKIQKAEAGITPQAYVDMGAVGVKELWQLLDISYDKIRITDD 60
99 DYHEKVAQVFERLLAODDIYLGEYSGMYSVDEEFPFESQLAE-VFRDEAGNVTGGIAP 157
61 ERKQVKEVFERLLKQDILVGEYSGMYSVDEEFPFESQLAE-VFRDEAGNVTGGIAP 118
158 -SGHEVWVSSESYFLRLSKYQDRLEFFKAHPEFTTPDGRNLMRLNIEQLDGLAVSRITFTTGWGVPSNPKHVVYVVI 216
119 DSGHEVELVKSSEYFFNFINKYTDRLLEFFYDENPFIQPPSRKNEMINNFIKPGLEDIAVSR 178

QY 217 RTTWTGVPSPKHHVYVWIDALLNATAGYAQDEHGFNKGTVFHVGVKDLR 276
DB 179 RTSFDMGVRVPSNPKHHVYVWIDALVNISSGLYSDDTELFNKPWADI-HLMAKEIVR 237
QY 277 FHSIYWPILLMMLDVKLPRDLIAHGWFMKDGKMSKSGNNVYVPEMLVERIGLDPLRYYL 336
DB 238 FHSIYWPILLMALDPLPKKVPFAHGWFMKDGKMSKSGNNVDPNVLDIDRYGLDAYRYL 297
QY 337 MRNLPVSGDGTFTTPEYVGRINYLEANDLGNLNRVTSMINKYVFDGOIPAYVSGVTEFDH 396
DB 298 MRELPFGSDGVTFPEAFVERTNYDLANDLGNLNRVTSMINKYVPHGSLPAYQPKHELDE 357
QY 397 VLAEVAEKSIADFRTHMEAVDYPRALEAVMTLISRTNKYIDETAPWVLDKDEALRDQLAS 456
DB 358 KMEAMALETVKSFNDNMESLQFSVALSTVMKFSISRTNKYIDETQWPVLAKDENQREMLGN 417
QY 457 VMSHWQASIRVVAHLIEPFFMETSRAVLTOIG-----LEEVSLENLSLADFFADVTVVA 511
DB 418 VMAHLVNERIFATLLQFFTHAPREIFKQINNNPDLHQLDLSLQOYGM--LSEAITVTE 475
QY 512 KGTFIFRLNMEEEIAYIKQMEGNKPAVEKEMWPDVEVLKLNKDEIKFDFDKVEIRVA 571
DB 476 KPTIFRPLDTEABIAIKESMQPPK-SIKQSDP-----GKEQIDIKDFDKVEIKAA 527
QY 572 EVKEVSKVSGDKLLQRLDAGDGEDROLISGLIAKYYPNEQELVGVKQVIVANLKPKROM 631
DB 528 TIIDAENVKSEKLLIKVEL-DNEQRIQVSGIAKPY-RPEDIIGKVAVVTNLKPAKLM 585
QY 632 KKYVSQGMILSAHDGKLTLLTVDPAVPNGSVI 664
DB 586 GQ-KSEGMLSAEKDGVTLISLSAIPNGAVI 617

RESULT 7

US-09-134-000C-4034
; Sequence 4034, Application US/09134000C
; Patent No. 6617156

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4034
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4034

Query Match 42.9%; Score 1492.5; DB 4; Length 493;
Best Local Similarity 60.1%; Pred. No. 2.7e-125;
Matches 283; Conservative 79; Mismatches 102; Indels 7; Gaps 6;

QY 199 NEMLRNTEPGLDLAVSRITFTGVVPSPNKHVVYVWIDALLNATAGYAQDEHGNF 258
DB 24 NEMLRNTEPGLDLAVSRITFTGVVPSPNKHVVYVWIDALLNATAGYAQDEHGNF 83
QY 259 DKFWNGTVFHVGVKDLRHSIYWPILLMVDVLPRLIAHGWFMKDGKMSKSGNNV 318
DB 84 QKYPANV-QVNGKEIVRFHTIYVFMIMALDPLPKKVPFGHGMLLKMGKMSKSGNNV 142
QY 319 YPEMLVRYGLDPLRYLIMRNLPGVSDGTFTPEYVGRINYLEANDLGNLNRVTSMINK 378
DB 143 YPEMLVRYGLDPLRYLIMRNLPGVSDGTFTPEYVGRINYLEANDLGNLNRVTSMINK 202
QY 379 YFDQIIPAYVGVTEFDHVLAEVAKSIADFRTHMEAVDYPRALEAVMTLISRTNKYIDE 438

DB 203 YCDGKVPAYASKVTFFDSELSTTAANVIGKVHEAMKMEFNATAIABITWLVSRANKYIDE 262
QY 439 TAPWVLDKDEALRDQLASVMSHWQASIRVVAHLIEPFFMETSRAVLTOIGLE-EVSSLEN 497
DB 263 TAPWVLAKEEERNEKRSVMIHLAELSRIVALLQPVMTETPGKIFEQGLDPEFTNMWEN 322
QY 498 LSLADFFADVTVVAKGTTFIFRLNMEEEIAYIKQMEGNKPAVEKE--WNPDEVELKLNK 555
DB 323 IHFGFPTDVTVTSGTTFIFRLE-ETEVTVIKQMSQSESAEEDIKWNPETTLVSTK 382
QY 556 D-EIKPEPDKVEIRVAEVEKVEKSGVSKLQFLDAGDGEDROLISGLIAKYYPNEQEL 614
DB 383 EQKIYDDFDKVELKVAEVIDCKVKYKADKLLQFLDAGDENHRLQISGLIAEFPDPAAL 442
QY 615 VGVKQVIVANLKPKROMKKYVSQGMILSAEH-DGKLTLLTVDPAVPNGSVI 664
DB 443 ICKKVIVANLKPRK-MRQIISQGMILSAESPEGLQIVEAPKEMPNGAGI 492

RESULT 8

US-08-451-715A-4
; Sequence 4, Application US/08451715A
; Patent No. 5801013

GENERAL INFORMATION:

; APPLICANT: Tao, Jianshi
; APPLICANT: Qui, Yan
; APPLICANT: Houtan, Fariba
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising Same
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,715A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-25
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-451-715A-4

Query Match 32.7%; Score 1137.5; DB 1; Length 648;
Best Local Similarity 38.9%; Pred. No. 3.2e-93;
Matches 266; Conservative 117; Mismatches 233; Indels 67; Gaps 16;

QY 8 ITTPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFFYLTGLDGHGKIQKAEAGIT 67
DB 6 ITTPIYVNDIPHIGHAVTTIADTLKKYVTLQGEVFFLTGTDBHGKIQEASLRNQS 65
QY 68 POAYVDGNVAGVKEWQLDLSIDYKFIETDDYHEKVAQVFERLLAODDIVLGSYGNV 127

Db 66 PRAYADISAIKFDQWDFNLDYDGFIRTTDSEHQKVCQNAPEIMFKEGDIYKGAISGY 125
 QY 128 SVSDEFFTESQLAEVFRDEAGNVGGIAPSG-HEVEWVSSESYELRLSKYQDRLEVEPK 186
 Db 126 CVSCESYCAISK-ADNTSKV-----LCPDCLRTALLEESYFFKUSAYEKPLUEFYA 178
 QY 187 AHPEFTTPDGRNLEMRNIEPGLEDLAVSRITFTWGVVPV---SNPKHVVVVWIDALLN 243
 Db 179 KNEPALPIYRKNB-VTSIEQGLDLSITRISFEGIPLPKMKNDPKHVVVVWIDALLN 237
 QY 244 YATAGYADQEHNGDFKFWNGTVFHVHVGDIIRFHSIYWPILLMMDLVKLPRLIAHGF 303
 Db 238 YASALGYLGNLKNMAHF--ECARHTVGVKDIIRFHAIFYWPAFLMSLNLPLFKQLCVHGM 295
 QY 304 VMKDGKWSKSGNVVPEMLVERYGDDPLRYILMRNLPVSGDTFTPDYVGRINVELAN 363
 Db 296 TIEGVZKSKSLGNVLDAQKLAMEYGEIELRYFLLEVPFGQDGEKSKALIERINANLNN 355
 QY 364 DLGNLLNRITVSMINKYFD-----GQIPAYVEGVTEPDHVLAEVAKSIADFHTHWEADV 418
 Db 356 DLGNLLNRLLGMKAKYFNYSLSKSAKITAYYSKELEKAOILDNANSFVPMQLH----- 409
 QY 419 PRALAVWTLISHTNKYIDETAPVWLDKDEALRDQLASVMSHWQASIRVVAHLIEPFNME 478
 Db 410 -KALELFNVYDFINKLIAKEEPWLVKNE-SEKLEALLSLIANALLOSSFLYAFMPK 467
 QY 479 TSRAVLTIQIGLE-----EVSSLENLSLADPPADVTVVAKGTPPIPRLNMBEEIAY 528
 Db 468 SAVKLASAFNTEITPNRYERFFKAKLQDWLQD-----TEPLFCMK----- 509
 QY 529 IKQMEGNPAVKEWNPDEVELK-----LNKBIKEPDEKVEIRVAEYKESYKVEG 581
 Db 510 --EKIEKTDKR-EKYPPEKAEKEKEKAPKQENYIGIEDFKVEIKVGLIKEAQRIEK 566
 QY 582 SDKLLQFRLDAGGEDRQILSGIYAKYVNEQELVGGKVOIVANLKPRKMKYVVSQGMIL 641
 Db 567 SNKLLBLKVDLGNRLRQIISGLALDYEP-SLVGQWCVVANLKPALMGE-MSEGMIL 624
 QY 642 SAHDGKLLTLLVDPVAVNGSVI 664
 Db 625 AVRDSNDNALISPTKEKIAGSLI 647

RESULT 9

US-08-584-226-2
 ; Sequence 2, Application US/08584226
 ; Patent No. 5798240
 ; GENERAL INFORMATION:
 ; APPLICANT: Martinis, Susan A.
 ; APPLICANT: Sasanfar, Mandana
 ; APPLICANT: Kim, Sunghoon
 ; APPLICANT: Lee, Sang Ho
 ; APPLICANT: Schimmel, Paul R.
 ; TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL METHIONYL-CRNA
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02173-4799
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/584,226
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/305,766
 FILING DATE: 13-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Brook, David B.
 REGISTRATION NUMBER: 22,592
 REFERENCE/DOCKET NUMBER: CP194-05Z
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 521 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-584-226-2

Query Match 29.7%; Score 1033; DB 1; Length 521;
 Best Local Similarity 41.0%; Pred. No. 5 5e-84;
 Matches 214; Conservative 94; Mismatches 200; Indels 14; Gaps 7;

QY 1 MSEKNFYITPIYPSGKHLIGSAYTTIACDLVARYKRLMGYDVYLTGLDEHGKIQOK 60
 Db 1 MEMKPYVTTTATAYPNAAPHVGHAYEYIATDARFRLDGYDVRFLTGTDEHGLVAQA 60
 QY 61 AEEAGITPOAYVDGMAGVKELWQLDISYDKFIRTTDDVHEKVAQVPERLLAQQDIYL 120
 Db 61 AAAGVPTAALARNNDVFORMQEALNISDFRITRTDADHHEASKELWRMSAAGDIYL 120
 QY 121 GEYSGWVSDEBFFBESQLAEVFRDEAGNVGTGIIAPSGHEVEWVSSESYFLLSKYQDR 180
 Db 121 DNTSGWVSVERDERFFVESETQLV-----DGTRLTVEITGTFVTWTEQTYFFRLSAYTDK 174
 QY 181 LVEFFKAPHFITPDGRNLEMRNIEPGLEDLAVSRITFTWGVVPSPKHVVVWIDA 240
 Db 175 LLAHYANPDIAPETRRNEVI-SFVSGGLDLSISRTSFDWGVQVPEHPDHHVYVWVA 233
 QY 241 LLNVATAGVAQDBHGNDFKFMNGTVFHMVGKDIILRFHSIYWPILLMMDLVKLPDLRIAH 300
 Db 234 LTNVLTGAGPDDTDSLEFRYWPADL-HMIAKDIIRFHAVYVWPAFLMSAGIELPRIFAH 292
 QY 301 GWFVMKDGKMSKGNVYVPEMLVERYGLDPLRYILMRNLPVSGDGTFTPDYVGRINYE 360
 Db 293 GFLHNRGEKMSKSVGNIVDPVALAEALGVQVRYFLLEVEYFPGQGSYSDEAIVTRINTD 352
 QY 361 LANDGLNLLNRITVSMINKYFDGQIPAYVEGVTEPDHVLAEVAKSIAADPHTHMEAVDYPR 420
 Db 353 LANELGNLARRSLSMVAKNLDGRVNPGE-FADADAALLATADGLLERYEGHFDQAQMH 411
 QY 421 ALEAVWTLISRTNKYIDETAPVWLDKDEALRDQ--LASVMHMQASIRVVAHLIEPFNME 478
 Db 412 ALEAILWMLGDANKYFSVQPPVWLKSESEADQARFRTTLYVTCEVVVRIAALLIQVMPE 471
 QY 479 TSRAVLTIQIG-LEEVSSLENLSLADPPADVTVVAKGTPPIPR 519
 Db 472 SAGKILDLQAPNQRSAFVAVGRLTGE--TALPPPTGVFPR 511

RESULT 10

US-09-543-681A-5403
 ; Sequence 5403, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 5403

; LENGTH: 683
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5403

Query Match 21.7%; Score 755; DB 4; Length 683;
Best Local Similarity 29.0%; Pred. No. 8e-59;
Matches 203; Conservative 137; Mismatches 276; Indels 84; Gaps 22;

QY 8 ITTPIYPSGKLHIGSAYTTIACDLVARYKRLMGYDVFYLTGLDEHGKIQOKAEAGIT 67
DB 18 VTCLPANGSIHLGHLEHIQADIVWVYQRMGEVHFICADDAHGTPIMLKAQOLGIT 77
QY 68 PQAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAVQVPERLLAODDYLGEVSGWY 127
DB 78 PEAMIEMSKEHQDQFAGNISYVHSTHSEENROLSTKIYALAKNKGHIKSKTISQY 137
QY 128 SVSDEE--FFTESQLAE-----VFRDEAG-----NVTGGIAPSGH 160
DB 138 ---DEEKGMPDPFRVKGTCPCCKAPQYGDNCVCGTSYSPTELINRSVSGSPVNR 194
QY 161 EVWVSEESYFLRLSKYQDLVEFFKAHPFITPDGRINEMLEN-----FIEGLEDLAYS 216
DB 195 ET-----EHYFFDLPAPSNLQEWIKS-----GALQEQVANKMQEWPDSGLQOWDIT 241
QY 217 RTTFTWGPVPSPKPHVYVWIDALLNYATGALGYAODEHG--NFDKFMN---GTVFHMV 270
DB 242 RDAPYGFGEIPDAPGKYFYVWLDAPICYMGSLNLCCKGDLSDFFWNKESKTDLYHFI 301
QY 271 GKDLRPHSTIYWPILLMLMDVKLPDLRLAHGFWMDGKMSKGNVYVPEMLVERYGLD 330
DB 302 GKDIVFHSLSFWPAVLEGSEYRKPTNLFVHGYYTVNGAKMSKRGTFITARAYLDHFDAD 361
QY 331 PLAYILMRNLVPVSGD--TFTPEYVGRINVELANDLGNILNRTVSMINKYFQGOIPAYVE 389
DB 362 CLRYVYAAKLSSRIDDLNLEDFVQVNSDIYKVVNLSRRTAGFISRRFDGKLAASLD 421
QY 390 GVTFEDRHLVAEAKSIADPHTHMEAVDYPRALAEVWTLISRTNKYIDETAPVWLKDDEA 449
DB 422 DAKLYQHFDV--NCEISQAQCFENR---FGKAIREIMALADEANRYDEKAPVWAKQEG 476
QY 450 LRQLASVMSHWQASIRVVAHLIEPP---NMWTSRAVL--TQIGLEEVSS--LENLSLADPP 504
DB 477 KEAELOAICTMGNLNFVLMYTKLPILPSLTERSEAPLQTLTWNALSQPLLNEHTKFK 536
QY 505 ADVTWVAKGTPIPRLNMEEIEAIYKEQMEGNKPAVE--KEMNPDEVELKLNKDEIKFED 562
DB 537 A-----LNRIMDXANAV-----EASKSTIAPVKEITGPLADAPI-QETIKFDD 581
QY 563 FDKVEIRVAEYKSVKVEGSDKLLQFRLDAGDGEDROILSGIAKYYPNEQELVKKQIV 622
DB 582 FAKIDMFIABIKQADFVDGSDKLLKLLDLDG-GETRQVFSGIRSAYPDPKVLGRLTVW 640
QY 623 ANLKPRLKMKKYVSQGMILSAEHGK-LTLLTVDPAVPNG 661
DB 641 ANLAPRK-MRFGISEGMVMAAGPGGEDIFLLSPDSGAKPG 679

RESULT 11
US-09-489-039A-8243
; Sequence 8243, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8243

; LENGTH: 716
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8243

Query Match 20.9%; Score 726; DB 4; Length 716;
Best Local Similarity 26.8%; Pred. No. 3.5e-56;
Matches 195; Conservative 129; Mismatches 281; Indels 122; Gaps 18;

QY 4 KNFVITPIYPSGKLHIGSAYTTIACDLVARYKRLMGYDVFYLTGLDEHGKIQOKAE 63
DB 45 KKILVTCLPANGSIHLGHLEHIQADIVWVYQRMGEVHFICADDAHGTPIMLKAQ 104
QY 64 AGITPQAYVDMGAVGVKELWQLLDISYDKFIRTTDDYHEKVAVQVPERLLAODDYLGEY 123
DB 105 LGITPEQIMIGSEHQDQFAGFIDISYDNHSTHSDENRELSELIYTRL----- 153
QY 124 SGWYSVSDERFFTESQLAEVFRDEAG----- 149
DB 154 -----KENGFIKRTTISQLYDPEKGMFLPDRFVKGTCPCCKSPQYGDNCVEVGATYSP 207
QY 150 -----NVTGGIAPSGHVEWVSEESYFLRLSKYQDLVEFFKAHPFITPDGRINEM 202
DB 208 TELDPSKSVSGATPVNRD-----SEHFFDLPFSFSEWLOAM-----TRSGALQEV 254
QY 203 RN-----FIEPGLLEDLAVSRITFTWGPVPSPKPHVYVWIDALLNYATGALGYAODEHG-- 256
DB 255 ANKMQEWPESGLQOWDISRDAPYGFGEIPNAPGKYFYVWLDAPICYMGSEFKNLCDXRGDT 314
QY 257 -NFDKFMN---NGTVFHMVMDKILLRHSIYWPILLMLMDVKLPDLRLAHGFWMDGKMS 311
DB 315 TSFDEYWKOKTABELYHFIGKDIYTFHSLFPAWLEGSNFRKPNLNFVHGYYTVNGAKMS 374
QY 312 KSKGNVVPPEMLVRYGLDPLRYLMLRNLVPVSGD--TFTPEYVGRINVELANDLGNILN 370
DB 375 KSRGTFIKASTLNLNHFADADSLRYIYVYAKLSRIDDLNLEDFVQVNSDIYKVVNLS 434
QY 371 RTVSMINKYFQGOI-----PAYVEGTFEDRHLVAEAKSIADPHTHMEAVDYPRALAE 424
DB 435 RNAGFISRRFDGKLAASLDAPALYKTFD-----AAESIGE-----AWDSREFGKAIRE 483
QY 425 VMTLISRTNKYIDETAPVWLKDDEALRDQLASVMSHWQASIRVVAHLIEPEFMETSRAVL 484
DB 484 IMALADVNRVYDGOAPVWAKQEGRDADLQAICTMGLNMFVRLMTWLKP-----VL 535
QY 485 TQIGLEEVSSLENLSLADFPADVTVVAKGTP---FFPRLNMBEELAYI---KEQMEGNKP 538
DB 536 POLAARAEAFNLSELSDAIQQPLLAHKVNPFKALYNRIEMKQVEALVEASKEEYKATAA 595
QY 539 AYEKEWNPDEVELKLNKDEIKFEDPKVEIRVAEYKSVKVEGSDKLLQFRLDAGDGEDR 598
DB 596 PVTGPLADDPQI-----ETITFDDFAKVDLRLVALIENAEFVEGSDKLLRLTLDLQ-GEKR 649
QY 599 QILSGIAKYYPNEQELVKKQIVANLKPRLKMKKYVSQGMILSAEHGK-LTLLTVDPA 657
DB 650 NVFSGIRSAYPDPQPLIGRLTVWANTAPRK-MRFGISEGMVMAAGPGGKIDIFLLSPDDG 708
QY 658 VPNGSVI 664
DB 709 AKPGQOV 715

RESULT 12
US-09-248-796A-16768
; Sequence 16768, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16768
; LENGTH: 402
; TYPE: PR1
; ORGANISM: Candida albicans
; US-09-248-796A-16768

Query Match 20.2%; Score 701.5; DB 4; Length 402;
Best Local Similarity 38.7%; Pred. No. 2.1e-54;
Matches 147; Conservative 77; Mismatches 141; Indels 15; Gaps 6;
QY 2 SEKNFVITPIYPSGKLHIGSAITTIACDVLARYKRLMGY-DVFLYTLGLDEHGQKIQQK 60
Db STKPFVITPIFYVNAAPHIGHLYSMLIADTRNKWEKLNPLKESFMTGTDEHGLKIQT 78
QY 61 AEEAGITPOAYVDMGAVGKELWQLLDISYDKFIRTTDDYHEKVAQVPERLLAQQDIYL 120
Db AEKLGLEPKVLVDKVSQNSFKAEQDVNYDRFIRTTDNHIELVRYFVNLMEKGFYIT 138
QY 121 GEYSGWYSDESBEFFTESQLAEFVDEAGNVGTGGIAPSGHEVWSEBSYFLRLSKYQDR 180
Db DTHSGWYS-DEFFETQIEEVKN--GKAVKISSETKNEVYQKETYFFKLSMFQEQ 196
QY 181 LVFFKAHPEFITPDGRLEMLNFTPEGLDLAVSR--TTTGVGVPSNPKHVVYWI 238
Db LIQFLKQNEPEFIKPHRYQFIKLEEDTKLPDLSISRPSRLKWSIEVPENDSTQKIYWF 256
QY 239 DALLNATATGVAQDEHGNFDRF-----WNGTVMVGVKDLRPHSIYWPILLMLDV 291
Db DALLNVLTKTPHGEVQDSKEVTPENSINPAT--HVIGKDIIRFKCIYWFILMAAGI 314
QY 292 KLPDRIAHGWFMKDGKMSKGNVYPEMLVERYGLDPLRYLMLRNLPVSGDGTFTPE 351
Db ELPKQVIVHSHMLCDGFKMSKLSGNVDPMEISEYGVDPVFFLVSNSIDDDCKFS-E 373
QY 352 DVYGRINVELANDLGNLAR 371
Db ELQQRSDAVLGKYNCLISR 393

RESULT 13
US-09-328-352-4765
; Sequence 4765, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4765
; LENGTH: 720
; TYPE: PR1
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-4765

Query Match 19.0%; Score 661; DB 4; Length 720;
Best Local Similarity 28.3%; Pred. No. 2.4e-50;
Matches 207; Conservative 115; Mismatches 296; Indels 114; Gaps 26;
QY 2 SEKNFVITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFLYTLGLDEHGQKIQQKA 61
Db SVRKILVTNALPYANGPIHGLHLYQADIWVRAMRANGHDVTVVCADDAHGTALMLRA 92
QY 62 ERAGITPOAYV-----DGMVGVKELWQLLDISYDKF-----IRTTDDYHE-- 102
Db EANGISPEEQIANVQKEHIRDFDGFVPHD-----YDTHSDANKARSTDIYIKNR 144

QY 103 -----KVAQVFE---RLLAQDDIYLG-----EYSGWYSVDSBEFFFTESQLAEFVRDE 147
Db EAGNIAVRPVTQLDFPEKGMFLSDRFIKGTCPKSGEDQYGDSCVCGTITYNATELLNPR 204
QY 148 AGNVGTGIIAPSGHEVWSEBSYFLRLSKYQDRSLVEFFKAHPEFITPDGRLEMLRN-- 204
Db --STLSGATP-----VE-KSSDHYFFKLPNFAELQKWTR-----DEGLPLSIAINKLD 250
QY 205 -FIEPGLLEDLAVSRITFTWGPVPSNPKHVVYVIMDALLNYATL-GYAQDEHG--NFDK 260
Db EWPEAGLADWDISRDAPYFGFEIPDAPNPKYFYVWDAPIGYMSSFENYIKTRKPDNLNFD 310
QY 261 FW-----NGTVFHVGVKDLRPHSIYWPILLMLDVKLPDLIAHGWFMKDGMSKGN 316
Db FWKDSQNEVYHFGKDIVYFHALFWPAMLEGANYRTFTGLFVNGELTVNGQVSKSRGT 370
QY 317 VVYPEMLVERYGLDPLRYLMLRNLPVSGDGTFTPEYVGRINVELANDLGNLNLRTVSM 375
Db FIKAEVTLQHLNPEYLEYFASKLSDKVEDSLNLDPFYQVNSDLVGKVVNIASRCAPF 430
QY 376 INKYFDGQIPAYVEGVTEFPHVLAEV-AEKSIADFTHTHEAVDYPRALEAVWTLIISRTNK 434
Db INSGFNNTLSS--TCAESDLVQSFIDAGDSIAAAY---EAREFSTAIRBIMALADRANQ 484
QY 435 YIDETAPWLVDKDEALRDQLASVMHQAASIRVVVAHLIEPFMMETSRVLTQIGLEEVSS 494
Db YIDKKWALAKGEGQBQVLDVCSGVINLFRQAVLAPVLPTLAQQVDFLKLESF-- 542
QY 495 LENSLADFPADVTV-----VAKGTTFPRLNMEESIAVI---KEQMEGNKP-----AVEK 542
Db -----DFESRQILVSHIEIAQFQPLMQRVDPKAVAAVMDASKESLIGAPAPQATKAAXK 595
QY 543 EWNPD-----EVELKLNKDEIKFDFDKVEIRVAEVEKSVKVGSDKLLQFRLDA 592
Db EKSAEKKAAPTAVGEABI-----IGIEDFLKVDLRVAQVVEAGTVEGSDKLLQUTLVD 649
QY 593 GDGEDRQILSGIARYPNEQELVGKVKQVIVANLKPRMMKKYVSQGM1-SAEHDGKLTLL 652
Db GEAEPRNVFSGIRSQYAPB-DLAKGLVVMVANLAPRK-MRFGISNGWVLAAGNEGIFII 707
QY 653 TVDPAVNGSVI 664
Db SPDSGAKPGDKV 719

RESULT 14
US-09-540-236-3507
; Sequence 3507, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3507
; LENGTH: 690
; TYPE: PR1
; ORGANISM: Moraxella catarrhalis
; US-09-540-236-3507

Query Match 18.2%; Score 633.5; DB 4; Length 690;
Best Local Similarity 28.3%; Pred. No. 6.6e-48;
Matches 204; Conservative 121; Mismatches 296; Indels 101; Gaps 26;
QY 2 SEKNFVITPIYPSGKLHIGSAITTIACDVLARYKRLMGYDVFLYTLGLDEHGQKIQQKA 61
Db TSRLKILVTSALPYANGPIHGLHLYVEIQDIWVRAMKAMGHNVTVCCADDAHGTALMLKA 69
QY 62 ERAGITPOAYVDMGAVGKELWQLLDISYDK-----FIRTTDDYHEKVAQVPERLLAQQD 117

Db	70	QONGITPBEQI-----ATVKASHKDFAGFLIDFDNTHS---THSSNQYFSOK 115
Qy	118	IYLG-EYSGWTSVSDDE--PFTESQLAEVRDEAGNVTGGIAPSGH-----EV-----162
Db	116	IYTLKONGYIFTKDVQELFOTEKGLFLADRFVKGTCPECDAPNOYGDNCNEVCSYTNAT 175
Qy	163	-----EWSVESYELRLSK-YQDRLVYFFKHAPEFITPDGRLEMLN-----FIEPGL 211
Db	176	EKNPYSTLUSNSPIIRSKHYFNLPOFDEFLRTWTKSQROQLSSYKNLKSFWPAGLA 235
Qy	212	DLAVSRITFTGCVYP-----SNPKHVYVYVTDALLNYATALGAQAQBHG-NPDKFNW---263
Db	236	PWDISKDAPYFGKIPDTPADBPDKYFYVWLDAVPVYMAFKNLCDRDTDMDFEYFNKDS 295
Qy	264	CTVPHMYCKDILRPHSIYWPILLMMLDVKLPDLRIAGHGFVKDQGWKSKGNVYVPEM 322
Db	296	QELYHFHIGKOLVYFHSUFWPAMLEGAGYRPTTAVNAHGFLTVNGEKOKSRGTFFIAET 355
Qy	323	LVERYGLDP--LRYVILMENLPVG-SDGTFPTDPPYVGRINYELANDLGNLNRITVSMINKY 379
Db	356	YLTH--LDEYLRVYFASKLSANVEDINLDEDFWQVNSDLVGKVNIASRAGSIHKN 413
Qy	380	FGQIPAYVEGVTEPDHVLAEVAKSIADPHTHMEAVDYPALAEVWTLISRTNKYIDET 439
Db	414	HOKUSDELTEPKLIDBICAGDEIAAA-----YEAREFAKVTLINACADKANEVIDTQ 468
Qy	440	APWLDK----DEALRDQLASV-MSHWQASIRVVAHL--TEPFMWETSRVLTOIGLEEV 492
Db	469	KPWALNKAGKFDQVKVCTVSLNIFH-----KLIIYAPILPKLSEQAREFLNIDSLEFD 523
Qy	493	SSLENI---SLADFPADYTVVAKGTPIFPRLNMBEEIAYIKEOMEGNKPAVEKEWNPDEV 549
Db	524	TRRHLLSHINPF-----TPLMQRIDKQVNMTHDSKEN-----LTQTSCTDNT 569
Qy	550	EIKLNKDE-----IKPEDFKVEIRVAEYKVEKSVGSKDLQLOFLDAGDGEDQILS 602
Db	570	KKTVLDDTOTAKSAVIGIDEFKAKVTMKVAHVVKQSVGEGADKLQLOFLDVGEEHTANVS 629
Qy	603	GIAKYYPNOELVGKKVQIVANLKPRKMKKYVSGMILSAEHGDKLTLLTVDPAPVNGS 662
Db	630	GIAKFTQPE-ELTDDKKVCTVNLAPRK-MKFGISEGMILSAEKQGVLTVITLDPDTPIGA 687
Qy	663	VI 664
Db	688	VL 689

RESULT 15
US-09-252-991A-17566
Sequence 17566, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.1136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17566
LENGTH: 680
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17566

Query Match 18.0%; Score 625; DB 4; Length 680;
Best Local Similarity 27.2%; Pred. No. 3.8e-47;
Matches 193; Conservative 130; Mismatches 302; Indels 84; Gaps 23;

Search completed: December 17, 2004, 22:01:06
Job time : 44 secs

This Page Blank (uspto)

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:56:12 ; Search time 147 Seconds

(without alignments)
1618.777 Million cell updates/sec

Title: US-10-606-607-2

Perfect score:

Sequence: 1 MSEKNFYITPTIYPSGKLH.....DGKLTLLTVDPAPVNGSVIG 665

Scoring table: BLOSUM62

scoring cable: 220300002
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Minimum	DB seq	Length:	0
Maximum	DB seq	Length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 100%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US07 PUBCOMB.pcp:
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB.pcp:**
3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pcp:**
4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pcp:**
5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pcp:**
6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pcp:
7: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pcp:**
8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pcp:
9: /cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB.pcp:
10: /cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB.pcp:
11: /cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB.pcp:
12: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pcp:
13: /cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB.pcp:
14: /cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB.pcp:
15: /cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB.pcp:
16: /cgn2_6/ptodata/2/pubpaa/US10D PUBCOMB.pcp:
17: /cgn2_6/ptodata/2/pubpaa/US10E NEW PUB.pcp:
18: /cgn2_6/ptodata/2/pubpaa/US11 NEW PUB.pcp:
19: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pcp:
20: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pcp:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3479	100.0	665	15	US-10-606-607-2
2	3446	99.1	679	9	US-09-815-242-13418
3	3446	99.1	679	15	US-10-282-122A-73893
4	2883.5	82.9	666	15	US-10-282-122A-74415
5	2883	82.9	669	15	US-10-282-122A-74425
6	2295	66.0	669	15	US-10-282-122A-57809
7	2288	65.8	677	9	US-09-815-242-10663
8	2288	65.8	677	15	US-10-282-122A-43475
9	2284.5	65.7	666	9	US-09-815-242-4933
10	2113	60.7	664	15	US-10-282-122A-60385
11	1949	56.0	559	15	US-10-282-122A-71524
12	1935.5	55.6	669	15	US-10-282-122A-46123
13	1934	55.6	557	9	US-09-815-242-12135
					Sequence 2, Appli
					Sequence 13418, A
					Sequence 73893, A
					Sequence 74415, A
					Sequence 74223, A
					Sequence 57809, A
					Sequence 10663, A
					Sequence 42475, A
					Sequence 4933, Ap
					Sequence 60385, A
					Sequence 71524, A
					Sequence 46123, A
					Sequence 12135, A

Query Match	100.0%	Score 3479;	DB 15;	Length 665;
-------------	--------	-------------	--------	-------------

Query Match	100.0%;	Score 3475; DE 15;	Length 6087
Best Local Similarity	100.0%;	Pred. No. 5e-263;	
Matches 665; Conservative	0; Mismatches 0;	Indels 0;	Gaps 0;

1 MSEKNEVITTPITVYPSGKI.HIGSAYTTIACDVLARYKRLMGYDVFYLTGLDEHGQKIQQX 60

1 MSEKNFILLILFLIFPSKMLTGLAIIACDVFENKRLICDVIICGLDGLK
1 MSEKNFVITPTIVVPSGKIHTGSAYTTIACDVIARYKSLNGYDVFFLTGLDEHGQKIQOK 60

Db5
I MSERKNFYLTTPAYYPFGKATHIGSAIIITACDVFARINAKLMSIDVPIHIGCEGEMHICQAWK:3

ALIGNMENTS

RESULT 1

```

US-10-606-607-2
; Sequence 2, Application US/10606607
; Publication No. US2004001421A1
; GENERAL INFORMATION:
; APPLICANT: LAWLOR, ELIZABETH JANE
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GM10253B-3C1D1
; CURRENT APPLICATION NUMBER: US/10/606,607
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: 10/187,641
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/025,189
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/432,692
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: 08/844,056
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: GB 9607999.1
; PRIOR FILING DATE: 1996-04-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-606-607-2

```

Sequence 43933, A
Sequence 13013, A
Sequence 17042, A
Sequence 5896, Ap
Sequence 52642, A
Sequence 51647, A
Sequence 53102, A
Sequence 46354, A
Sequence 54439, A
Sequence 76747, A
Sequence 1554, A
Sequence 6988, Ap
Sequence 5970, Ap
Sequence 5890, Ap
Sequence 5709, A
Sequence 62338, A
Sequence 62166, A
Sequence 63675, A
Sequence 64852, A
Sequence 13121, A
Sequence 64107, A
Sequence 63402, A
Sequence 10360, A
Sequence 163558, A
Sequence 7843, Ap
Sequence 68671, A
Sequence 10210, A
Sequence 5660, A
Sequence 75322, A
Sequence 76244, A
Sequence 13925, A
Sequence 4773, Ap
Sequence 2, Appli

```

61 AEEAGITPQAVYDGMAGVVKELMQLLDSYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
61 AEEAGITPQAVYDGMAGVVKELMQLLDSYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
121 GEYSGWYSDDEBFFTESQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
121 GEYSGWYSDDEBFFTESQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPXHVYVWIDA 240
181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPXHVYVWIDA 240
241 LLNYATGALGYAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH 300
241 LLNYATGALGYAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH 300
301 GWFVKDGMKSKGNVYVPEMLVERVGLDPLRYILMRNLVPGSDGTFPTPEDVGRINYE 360
301 GWFVKDGMKSKGNVYVPEMLVERVGLDPLRYILMRNLVPGSDGTFPTPEDVGRINYE 360
361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYP 420
361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYP 420
421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVEGSKLLQFLDAGDGEDROI 600
541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVEGSKLLQFLDAGDGEDROI 600
601 LSGIAYKYPNEQELVGKQVIVANLAKPRKMKYVSGMILSAEHGDKLTLLTVDPAVFN 660
601 LSGIAYKYPNEQELVGKQVIVANLAKPRKMKYVSGMILSAEHGDKLTLLTVDPAVFN 660
661 GSVIG 665
661 GSVIG 665

```

RESULT 2

```

US-09-815-242-13418
; Sequence 13418, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13418

Query Match      99.1%; Score 3446; DB 9; Length 679;
Best Local Similarity 99.2%; Pred. No. 2e-260;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSEKNFYITPIYVPSGKLGHSAYTTIACDVLARYKELMGYDVFYLTGLDEHGKIQOK 60
Db 15 MSEKNFYITPIYVPSGKLGHSAYTTIACDVLARYKELMGYDVFYLTGLDEHGKIQOK 74
Qy 61 AEEAGITPQAVYDGMAGVVKELMQLLDSYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
Db 75 AEEAGITPQAVYDGMAGVVKELMQLLDSYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 134
Qy 121 GEYSGWYSDDEBFFTESQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 180
Db 135 GEYSGWYSDDEBFFTESQLAEVPRDEAGNVTGGIAPSGHEVWSEESYFLRLSKYQDR 194
Qy 181 LVEFFKAHPFFITPDGRLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPXHVYVWIDA 240
Db 195 LVEFFKAHPFFITPDGRLNEMLNFIPEGLEDLAVSRITFTWGPVPSNPXHVYVWIDA 254
Qy 241 LLNYATGALGYAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH 300
Db 255 LLNYATGALGYAQDEHGNFDFWNGTVFHMVGKDIILRFHSIYWPILLMLLDVLPDLIAH 314
Qy 301 GWFVKDGMKSKGNVYVPEMLVERVGLDPLRYILMRNLVPGSDGTFPTPEDVGRINYE 360
Db 315 GWFVKDGMKSKGNVYVPEMLVERVGLDPLRYILMRNLVPGSDGTFPTPEDVGRINYE 374
Qy 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYP 420
Db 375 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTEFDFHVAEVAEKSIADPHTHMEAVDYP 434
Qy 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHQASIRVVAHLIEPFNMETS 494
Qy 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 554
Qy 541 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVEGSKLLQFLDAGDGEDROI 600
Db 555 EKWNPDDELKLNKDEIKFEDPKVEIRVAEYKEVSKVEGSKLLQFLDAGDGEDROI 614
Qy 601 LSGIAYKYPNEQELVGKQVIVANLAKPRKMKYVSGMILSAEHGDKLTLLTVDPAVFN 660
Db 615 LSGIAYKYPNEQELVGKQVIVANLAKPRKMKYVSGMILSAEHGDKLTLLTVDPAVFN 674
Qy 661 GSVIG 665
Db 675 GSVIG 679

```

RESULT 3

```

US-10-282-122A-73893
; Sequence 73893, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 73893
LENGTH: 679
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73893

Query Match 99.1%; Score 3446; DB 15; Length 679;
Best Local Similarity 99.2%; Pred. No. 2e-260;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSEKNFYITPIYPSGKLHGSAYTTIACDVLARYKRLMGYDVFLTGLDEHGQIKQ 60
DB 15 MSEKNFYITPIYPSGKLHGSAYTTIACDVLARYKRLMGYDVFLTGLDEHGQIKQ 74
QY 61 ABEAGITPQAYVDGMAGVKELWQLLDSYDKFIRITDDYHEKVAQVFERLLAQDDIYL 120
DB 75 ABEAGITPQAYVDGMAGVKELWQLLDSYDKFIRITDDYHEKVAQVFERLLAQDDIYL 134
QY 121 GEYSGWYSVDEBFTESQLAEVRDEAGNVTGCIAPSGHEVWVSEESYFLRLSKYQDR 180
DB 135 GEYSGWYSVDEBFTESQLAEVRDEAGNVTGCIAPSGHEVWVSEESYFLRLSKYQDR 194
QY 181 LVBEFKAHPEFITPDGRNLEMRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 240
DB 195 LVBEFKAHPEFITPDGRNLEMRNFIEPGLDLAVSRITFTWGPVPSNPKHVVYVWIDA 254
QY 241 LLNATYALGVAQDEHGFDFWNGTVFHMVYKQILRPHSIYWPILLMMLDVKLPDRLIAH 300
DB 255 LLNATYALGVAQDEHGFDFWNGTVFHMVYKQILRPHSIYWPILLMMLDVKLPDRLIAH 314
QY 301 GWFVWKDGKSKSGNVVYEMLVRYGLDPLRYLNRNLPVSGDGTFTFEDVVGRIINYE 360
DB 315 GWFVWKDGKSKSGNVVYEMLVRYGLDPLRYLNRNLPVSGDGTFTFEDVVGRIINYE 374
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTFEDHVLAEVAKSIADFTHTHEAVDYPR 420
DB 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTFEDHVLAEVAKSIADFTHTHEAVDYPR 434

QY 421 ALEAVTTLISRTNKYIDTAPFVLDKDEALRDQALASVMHQAASIRVVVAHLIEPMMETS 480
DB 435 ALEAVTTLISRTNKYIDTAPFVLDKDEALRDQALASVMHQAASIRVVVAHLIEPMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLANNEBEIAYIKQMEGNKPAV 540
DB 495 RAVLTQIGLEEVSSLENLSLADFPADVTVAAGTPIPPRLANNEBEIAYIKQMEGNKPAV 554
QY 541 EKEMNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSKLLQFRLDAGDGEDRQI 600
DB 555 EKEMNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVEGSKLLQFRLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKQVQIVANLKPRKMMKKYVSQGMILSAHDGKLTLLTVDPVAPN 660
DB 615 LSGIAKYYPNEQELVGKQVQIVANLKPRKMMKKYVSQGMILSAHDGKLTLLTVDPVAPN 674
QY 661 GSVIG 665
DB 675 GSVIG 679

RESULT 4
US-10-282-122A-74415
Sequence 74415, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74415
LENGTH: 666
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-282-122A-74415

Query Match 82.9%; Score 2883.5; DB 15; Length 666;

Best Local Similarity 81.9%; Pred. No. 1.9e-216;
Matches 546; Conservative 54; Mismatches 64; Indels 3; Gaps 2;
QY 1 MGEKNFYITPIYPSGKLHIGSAVTTIACDVLAARYKRLMGVDVFLYGLDEHGKIQK 60
Db 1 MMKPPYITPIYPSGKLHIGSAVTTIACDVLAARYKRLMGHEVFLYGLDEHGKIQK 60
QY 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIITDDYHKKVAQVFERLLAQDDIYL 120
Db 61 ABEAGITPQAYVDGMAVGKELWQLLDISYDKFIITDDYHKKVAQVFERLLAQDDIYL 120
QY 121 GEYSGWYSVDEEFTTESQLAEVDFDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 180
Db 121 GEYSGWYSVDEEFTTESQLAEVDFDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 181 LVEFFKAHPEFITPQGRNLMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWIDA 240
Db 181 LVAFFEREPDFIQDGRNLMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWIDA 240
QY 241 LNNYATGALGYQDDEHGNDFKFWGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDLIAH 300
Db 241 LNNYATGALGYQDDEHGNDFKFWGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDLIAH 300
QY 301 GWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
Db 301 GWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
QY 361 LANDLGNLNRVTSMINKYDFQCIYAYVEGVTPEFHLVAEVAEKSIADPHTHMEAVDYP 420
Db 361 LANDLGNLNRVTSMINKYDFQCIYAYVEGVTPEFHLVAEVAEKSIADPHTHMEAVDYP 420
QY 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQSLASVMSHQASIRVVAHLIEPFWMET 480
Db 421 ALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQSLASVMSHQASIRVVAHLIEPFWMET 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPLNMEEBIAIKQMEGNKPAV 540
Db 481 AATMAQLGEPVSDLSLALADFPADVTVVAKGTPIFPLNMEEBIAIKQMEGNKPAV 539
QY 541 --EKENPDEVELKLNKDEIKFDFDFAVEIRVAEVEKSVKSGSKLLQFLRDLADGD 598
Db 540 SOEKWVPEVALKSEKQVITETFDVAEIRVAEVEKSVKSGSKLLRFRVDAGDGD 599
QY 599 QILSGIAKYPNEQELVGGKQIIVANLKPQWKKYISQGMILSAEHGDKTLFLVDPAY 658
Db 600 QILSGIAKYPNEQELVGGKQIIVANLKPQWKKYISQGMILSAEHGDKTLFLVDPAY 659
QY 659 PNGSVIG 665
Db 660 PNGSIIG 666

RESULT 5
US-10-282-122A-72423
Sequence 72423, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Chlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72423
LENGTH: 669
TYPE: PRT
ORGANISM: Streptococcus mutans
US-10-282-122A-72423
Query Match 82.9%; Score 2883; DB 15; Length 669;
Best Local Similarity 82.4%; Pred. No. 2.1e-216;
Matches 553; Conservative 46; Mismatches 54; Indels 8; Gaps 5;

QY 1 MBEKNFYITPIYPSGKLHIGSAVTTIACDVLAARYKRLMGVDVFLYGLDEHGKIQK 59
Db 1 MTEKQFYITPIYPSGKLHIGSAVTTIACDVLAARYKRLMGVDVFLYGLDEHGKIQK 60
QY 60 KASEAGITPQAYVDGMAVGKELWQLLDISYDKFIITDDYHKKVAQVFERLLAQDDIY 119
Db 61 KSEAGITPQAYVDGMAVGKELWQLLDISYDKFIITDDYHKKVAQVFERLLAQDDIY 120
QY 120 LGYSWYSVDEEFTTESQLAEVDFDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 179
Db 121 LGYSWYSVDEEFTTESQLAEVDFDEAGNVGTGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 180 RLVEFFKAHPEFITPQGRNLMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWID 239
Db 181 RLVDFFHAHPEFITPQGRNLMRNFIPEGLEDLAVSRITFTWGVVPVSPNPKHVVTWID 240
QY 240 ALLNYATGALGYQDDEHGNDFKFWGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDLIA 299
Db 241 ALLNYATGALGYQDDEHGNDFKFWGTVFHMVGKDILRFHSIYWPILLMMLDVKLPDLIA 300
QY 300 HGWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 359
Db 301 HGWFVMDGKWSKSGNVVYPEMLVEREGLDPLRYLLMRSLPVGSDGTFPTPDYVGRINYE 360
QY 360 ELANDLGNLNRVTSMINKYDFQCIYAYVEGVTPEFHLVAEVAEKSIADPHTHMEAVDYP 419
Db 361 ELANDLGNLNRVTSMINKYDFQCIYAYVEGVTPEFHLVAEVAEKSIADPHTHMEAVDYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQSLASVMSHQASIRVVAHLIEPFWMET 479
Db 421 RALEAVWTLISRTNKYIDETAPWVLDKDEALRDLQSLASVMSHQASIRVVAHLIEPFWMET 480
QY 480 SRAVLTOIGLEEVSSLENLSLADFPADVTVVAKGTPIFPLNMEEBIAIKQMEGNKPA 539
Db 481 SNAIMEQLGAGQFDLEGLAGLPGVVKVIAKGTPIFPLNMEEBIAIKQMEGNKPA 539
QY 540 V---EKENPDEVELKLNKDEIKFDFDFAVEIRVAEVEKSVKSGSKLLQFLRDLADGD 595
Db 540 IQAEEKEWNPPTVELKNEKKAIFDFFDLKLEIRVAEVEKSVKSGSKLLKFLRLADGD 599

QY 596 EDQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGMILSASH-DGKLTLLTV 654
DB 600 EDQILSGIAKYPNEQELVKKVQIVANLKPRKMM-KLVSGMILSASHADGKLTLLTV 658
QY 655 DPAVPSNGSVIG 665
DB 659 DSSVPNGSQIG 669

RESULT 6
US-10-282-122A-57809
; Sequence 57809, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57809
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57809

Query Match 66.0%; Score 2295; DB 15; Length 669;
Best Local Similarity 65.2%; Pred. No. 2e-170;
Matches 438; Conservative 93; Mismatches 129; Indels 12; Gaps 9;

QY 1 MSEK-NEYITPIYVPSGKLHGSAYTTIACDVLARYKRLMGVDVFLTGLDGHGKIQQ 59
DB 1 MAEKETFIITPIYPSGKLHGSAYTTIACDVLARYKRLMGVDVFLTGLDGHGKIQET 60
QY 60 KAEKAGITPQAYVDMGAVGKELMQLLDISYDKFIRTTDDYHKKVAVQPERLLAODDIY 119
DB 61 KAEELGVKPCQYDKMAADVKKLWKTLDISYDKFIRTTDDYHKKVAVQKIFDRLLQGGDIY 120
QY 120 LGEYSGHYSDSEFFFTESQLAEVPRDEAGVTGGIAPSGHEVWSEERSYFRLSKYQD 179

DB 121 LGEYSGHYSDSEFFFTETQLAEVYKDDGNGVIGKAPSGHEVWSEERSYFRLSKYAD 180
QY 180 RLVEFFKRAHPFITPDGLNEMLNFEPLGLEDAVSRRTFTWGVVPSPKHYVYVWID 239
DB 181 RLLEYYEHPFITPDGLNEMLNFEPLGLEDAVSRRTFTWGVVPSPKHYVYVWID 240
QY 240 ALLNAYATAGYAQDEHGNFDFXWNGTVPHVMYKDIILRFHSIYWPILLMMLDKVLDRLLA 299
DB 241 ALSNVITAGYSEDESLEFKYWPADV-QMVGKEIVRFHTIYWIIMLMALDPLPKVFG 299
QY 300 HGWFVKDKGKSKSGNVVYPEMLVRYGLDPLAYLYMLNLPVGSDDGTFTPEDVVGRTNY 359
DB 300 HGWLLMKDKGKSKSGNVVYPEMLVRYGLDPLAYLYMLNLPVGSDDGTFTPEDVVGRTNY 359
QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGVTDFDHLVAEVAEKSIADFEHFMHMEADYD 419
DB 360 DIANDLGNLNRVTSMINKYFDGQIPAYVEGVTDFDHLVAEVAEKSIADFEHFMHMEADYD 419
QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDBALRDQASVMHSHWQASIRVAHLIEPFMMET 479
DB 420 TALAEIWLISRTNKYIDETAPWVLDKDBALRDQASVMHSHWQASIRVAHLIEPFMMET 479
QY 480 SRVLTQIGLB-EVSSLNLSLADFPADVTVVAKGTPIFFRLNMEEETIAYIKEQM-EQ-- 535
DB 480 PEKFNQGLDSETMNLEGLHFGFPSTGKVKVAKGTPIFFRLNMEEETIAYIKEQM-EQ-- 535
QY 536 -NKPAVEKMPNDEVELKLNKD-BIKFEDFDKXIRVAEVAEKSEKVEGSKLQFLRLDAG 593
DB 540 TNEDTV--KWDPEETELVSTKEKQIKFDFVEKVELKVAEINCKQVEGADKLLQFLRLDAG 597
QY 594 DCEBQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGMILSASH-HDGKLTLL 652
DB 598 DSQDRQILSGIAKYPNEQELVKKVQIVANLKPRKMMKKYVSGMILSASH-HDGKLTLL 652
QY 653 TVDPAVPSNGSVI 664
DB 657 BAPKSMPSNGSEI 668

RESULT 7
US-09-815-242-10663
; Sequence 10663, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10663

; LENGTH: 677

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-815-242-10663

Query Match 65.8%; Score 2288; DB 9; Length 677;

Best Local Similarity 63.9%; Pred. No. 7.3e-170;

Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPTIYPSGKLHGSAYTTTACDVLARYKELMGYDYFYLTLGDEHGQKIQ 59
 DB 9 MSQKETFYITPTIYPSGKLHGSAYTTTACDAMARYKELMGDFVPLTGVDEHGQKIEK 68
 QY 60 KAEAGITPOAYVDGMAVGKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
 DB 69 KAAELNVTPEYVDKAAADVQKWLKTLDISYDKFIRTTDDYHMAAQQIFDLRLVSGDIY 128
 QY 120 LGEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGTAPSGCHEVWVSESYFLRLSKYQD 179
 DB 129 LGEYEGWYSVDSDEFFTESQLAEVYRDDGNNVIGGKAPSGCHEVELYKEESYFFRMSKYAD 188
 QY 180 RLVEFFKAPHEFTTPDGLNLMELNFIETPGLEDLAVSRITFTWGVPSNPKHVYVWID 239
 DB 189 RLVOYEEHPEFTIOPESRKNEMLNFIKPGLEDLAVSRITFTWGVPSNPKHVYVWID 248
 QY 240 ALLNYATALGYAODEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
 DB 249 ALSNYITALGYGSEDDSLFQKYPANV-QMVGKEIVRFHTIYWPIMLMALDPLPKKVF 307
 QY 300 HGWFMKDGKMSKSGNVVYPEMLVERYGIDPLRYVLMNLPVSGDGTTPEDYVGRIN 359
 DB 308 HGWLLMKDGKMSKSGNVVYPEMLVERYGIDALRYLLRAIPFGSDGVFTPEDFVSR 367
 QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVGEVTEFDHVLAEVAEKSIADPHTMEAVDYP 419
 DB 368 DLANDLGNLNRVTSMINKYCDGKVPAYASKVTPFDSLSSTTAANVIGKYHEAMEKMEFN 427
 QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALDOLASVMSHQASIRVVAHLIEPMMET 479
 DB 428 TATAEITWLVSRANKYIDETAPWVLAKEEKEKNELESVMHIAESLRIVAILLOPVMTET 487
 QY 480 SRAVLITQIGLE-EVSSLENLSLADFPADVTVVAKGTPIPPRLNMBEEIAYIKEQMEGNKP 538
 DB 488 PGKIPFQGLDPTNMKNENHFEFTDVTIVTSKGTPIPFRLIEITEVTYIQQKMSQSES 547
 QY 539 AVEKE--WNPDEVELKIND-EIKFEDFDKVEIRVAEVEKVEGSKLLOFRILDAGDG 595
 DB 548 ATBEDIKWNPDEETLVSTKQIKYDDFDKVELKVAEVIDCKKVKGADKLLOFRILDAGDE 607
 QY 596 EDQIISLGIAKYPNBOELVKKVQIVANLKPDKMMKVVSGMILSABH-DGKLTLLTV 654
 DB 608 NHRQIISLGIAEFPDPAALIGKXVIVANLKPDK-MRQIISGMILSABSPGKLIQIVEA 666
 QY 655 DPAPVNGSVI 664
 DB 667 PKEMPNGAGI 676

RESULT 8

US-10-282-122A-42475

; Sequence 42475, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A US/10/282,122A

; CURRENT APPLICATION NUMBER: 2003-02-20

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: 60/269,308

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 42475

; LENGTH: 677

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-10-282-122A-42475

Query Match 65.8%; Score 2288; DB 15; Length 677;

Best Local Similarity 63.9%; Pred. No. 7.3e-170;

Matches 428; Conservative 104; Mismatches 130; Indels 8; Gaps 7;

QY 1 MSEK-NFYITPTIYPSGKLHGSAYTTTACDVLARYKELMGYDYFYLTLGDEHGQKIQ 59
 DB 9 MSQKETFYITPTIYPSGKLHGSAYTTTACDAMARYKELMGDFVPLTGVDEHGQKIEK 68
 QY 60 KAEAGITPOAYVDGMAVGKELWQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
 DB 69 KAAELNVTPEYVDKAAADVQKWLKTLDISYDKFIRTTDDYHMAAQQIFDLRLVSGDIY 128
 QY 120 LGEYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGTAPSGCHEVWVSESYFLRLSKYQD 179
 DB 129 LGEYEGWYSVDSDEFFTESQLAEVYRDDGNNVIGGKAPSGCHEVELYKEESYFFRMSKYAD 188
 QY 180 RLVEFFKAPHEFTTPDGLNLMELNFIETPGLEDLAVSRITFTWGVPSNPKHVYVWID 239
 DB 189 RLVOYEEHPEFTIOPESRKNEMLNFIKPGLEDLAVSRITFTWGVPSNPKHVYVWID 248
 QY 240 ALLNYATALGYAODEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIA 299
 DB 249 ALSNYITALGYGSEDDSLFQKYPANV-QMVGKEIVRFHTIYWPIMLMALDPLPKKVF 307
 QY 300 HGWFMKDGKMSKSGNVVYPEMLVERYGIDPLRYVLMNLPVSGDGTTPEDYVGRIN 359
 DB 308 HGWLLMKDGKMSKSGNVVYPEMLVERYGIDALRYLLRAIPFGSDGVFTPEDFVSR 367
 QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVGEVTEFDHVLAEVAEKSIADPHTMEAVDYP 419
 DB 368 DLANDLGNLNRVTSMINKYCDGKVPAYASKVTPFDSLSSTTAANVIGKYHEAMEKMEFN 427
 QY 420 RALEAVWTLISRTNKYIDETAPWVLDKDEALDOLASVMSHQASIRVVAHLIEPMMET 479
 DB 428 TATAEITWLVSRANKYIDETAPWVLAKEEKEKNELESVMHIAESLRIVAILLOPVMTET 487
 QY 480 SRAVLITQIGLE-EVSSLENLSLADFPADVTVVAKGTPIPPRLNMBEEIAYIKEQMEGNKP 538

```
Db 488 PGKIFQGLDPTMTNMENIHFGFTDVTVTSKGPPIFRLEIEFTEVYIYOKMSQSES 547
Qy 539 AVEKE--WNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKSGSKLLQFRILDAGDG 595
Db 548 ATEEDIKNPEETTLVSTKEKQIKYDDFDKVELKVAEVIDCKVKVGADKLLQFRILDAGDE 607
Qy 596 EDRQILSGTAKYYPNEQELVGKQVQIVANLKPRKMMKKTVSQGMILSAEH-DGKLTLLTV 654
Db 608 NHRQILSGTAEFPDPAALIGKVKVIVANLKPRK-MRQGISQGMILSAESPEGKQIIVEA 666
Qy 655 DPAVNGSVI 664
Db 667 PKMPNGAGI 676

RESULT 9
US-09-815-242-4933
; Sequence 4933, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4933
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4933

Query Match 65.7%; Score 2284.5; DB 9; Length 666;
Best Local Similarity 63.7%; Pred. No. 1.3e-169;
Matches 425; Conservative 105; Mismatches 130; Indels 7; Gaps 6;

Qy 3 EKFPYITTPYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFYLTGLDGHGKIOQKAE 62
Db 2 KETFYITTPYPSGKLHIGNSYTTIACDARVYKRLMGDFVYLTGLDGHGKIEKAA 61
Qy 63 EAGITPAYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDDIYLGE 122
Db 62 ELNVTPEYVDKMAADVQLWKTLDISYDKFIRTTDDYHMAAQVQIFDRLVEQGDIVLGE 121
Qy 123 YSGWYSVDEBFFTESQLAEPREAGNVGGTAPSGHEVETWSEESYFRLSKYODRLV 182
Db 122 YEGWYSVDEBFFETQLAEYRDEGNVIGGKAPSGHEVELVKESYFFRMSKYADRLV 181
Qy 183 EFFKAHPEFITPDGRNLNMLRNFIEPGLDLAVSRRTTFTWGVFVPSNPKHVVVWIDALL 242
```

```
Db 182 QYEHFEFTQPESEKKNMLNNFKPGLDLAVSRRTTFSNGIPKNDPKHVVVWIDALS 241
Qy 243 NYATALGYAQDEHGNFKFNGTTFVHMVGVKDIIRPHSIYWPILLMLLDVLPDRLIAHG 302
Db 242 NYITALGYSSDDSLFQKYNPANY-QMVGEIVAFHTIYWFIMLMALDLPKPKVFGHW 300
Qy 303 FVMKDGKMSKSGNVVPEMLVERYGLDPLRYVLMRNLPGVSDGTFTPTPEDYVGRINYELA 362
Db 301 LMKDGKMSKSGNVVPEMLVERYGLDPLRYVLMRNLPGVSDGTFTPTPEDYVGRINYELA 360
Qy 363 NDLGNLNRVTMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSIADPHTHMEAVDYPRAL 422
Db 361 NDLGNLNRVTMINKYCDGKVPAYASKVTPFDSELSTTAANVIGKYHEAMEKMEFTAI 420
Qy 423 EAVMTLSRTNKYIDETAPWLDKDEALRDOLASVMSHWQASIRVAHLTEPPMETSRA 482
Db 421 AEIWTLSRANKYIDETAPWVLAKEEKRNELESVHLHAESEIRVAILQPVMTETPGK 480
Qy 483 VLTQIGLE-EVSSLENLSLADFPADVTTVAKGTPIFRLNMEEFIAIYKMEGNKPAVE 541
Db 481 IFEQLGLDPTMTNMENIHFGFTDVTVTSKGPPIFRLEIEFTEVYIYOKMSQSESATE 540
Qy 542 KE--WNPDEVELKLNKD-EIKFEDFDKVEIRVAEVEKSVKSGSKLLQFRILDAGDEDR 598
Db 541 EDIKWNPBEETTLVSTKEKQIKYDDFDKVELKVAEVIDCKVKVGADKLLQFRILDAGDENHR 600
Qy 599 OILSGIAKYYPNEQELVGKQVQIVANLKPRKMMKKTVSQGMILSAEH-DGKLTLLTVDDPA 657
Db 601 QILSGIAEFPDPAALIGKVKVIVANLKPRK-MRQGISQGMILSAESPEGKQIIVEAPKE 659
Qy 658 VPNGSVI 664
Db 660 MPNGAGI 666

RESULT 10
US-10-282-122A-60385
; Sequence 60385, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60385
; LENGTH: 664
; TYPE: PR1
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60385

Query Match
  60.7%; Score 2113; DB 15; Length 664;
  Best Local Similarity 60.7%; Pred. No. 3.4e-156;
  Matches 407; Conservative 101; Mismatches 146; Indels 18; Gaps 8;

Qy 3 EKN-FYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVFLYLTGDEHGQKIQKA 61
Db 4 EKNFVITPIYPSGKAHIGAYTTVAGDARVYKLGKGYDVFLYLTGDEHGQKIQKA 63

Qy 62 BEAGITPQAYVDMGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 121
Db 64 KERGISEQYVDEIARGFQELWKKLEISNTDFIRTTQDRHKTSEVKFEQLLEQGGIYL 123

Qy 122 EYSGWYSVDEBFFTESQLAEVFRDEAGNVGTGIAPSGHEVWSEESYFLRLSKYQDL 181
Db 124 EYEGWYSVDEEYFTQLEEVYKDENGKVIQKAPSGNEVELVKEESYFFKMSKYADRL 183

Qy 182 VEFFKAHPFETPDGRLNEMLRNFIPLGLEDLAVSRTTFTWGPVPSNPKHVVYWDAL 241
Db 184 VEYNHSPFETPLPSKNEMINNFIPGLLEDLAVSRTTFTWGIKVPKPHVVYWDAL 243

Qy 242 LNAVATAGYAOEHGDFKFWGTWFMVYKGLILPHSYWELLMLMDVLDKLPDLRIAG 301
Db 244 SNVITAGVNTDNTKQYWPADV-QVIGKIVRPHITVYFIMLMALDLPFGVFGHG 302

Qy 302 WFMKDGKMSKSGNVVYPEMLVRYGLDPLRYLMRLNLPVSGDGTFTPEDYVGRINYL 361
Db 303 WILMKDGKMSKSGNVVDPYMLIDRYGLDALRYLLREVPFGSDGLFTPEDVDRVNFDL 362

Qy 362 ANDLGNLNRVSMINKYFDGQIPAYVEGVTEFDVLAEVAEKSTADFTHTMEAVDIPRA 421
Db 363 ANDLGNLNRVAMINKYFDGQIPAYVQGNVTEFDQILVDKNNVYKEVSGMDHMQFSA 422

Qy 422 LEAVMTLISRTNKYIDETAPWVLDDKDEALRDOLASVMSHWQASIRVAHLIEPFWMTSR 481
Db 423 LNLWLSLISRTNKYIDETAPWALAKDEKRTLASVMTHLAENLIIAVLQPLTRTPG 482

Qy 482 AVUTQGLSEVSLENL-----SLADPPAVTVVAKGTPIPPRLNMBEELIAYKEOME 534
Db 483 EIPLOLGLQES---ENLKKWDSIYGGEIPAGTTVVKKGTPIPPRLDAEVEVYIQDEM 538

Qy 535 GKNPAVEKSWNPEVELKLNKDEIKFEDDKVIRVAEYKESYKVEGSKLLQRLDAGD 594
Db 539 GSAPAPAEB--VAEVB-ALETQIGIEDFKIDLRVAEYKQVDRKKADKLLCFOLDGE 595

Qy 595 GEDRQILSGIAYYPNEQELGKVKQIVANLPRKMKKYVSQGMILSAEHDGKLTITV 654
Db 596 GKLRQVLSGIABFYQPE-ELIGKQVIVSNLKPVK-LRGLMSEGMLSGDKGLSVIEA 653

Qy 655 DEAVPNGSVI 664
Db 654 SSALPNGAKV 663
```

RESULT 11

```
US-10-282-122A-71524
; Sequence 71524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
```

```
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Onlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71524
; LENGTH: 659
; TYPE: PR1
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71524
```

```
Query Match 56.0%; Score 1949; DB 15; Length 659;
Best Local Similarity 57.3%; Pred. No. 2.3e-143;
Matches 383; Conservative 103; Mismatches 168; Indels 14; Gaps 9;

Qy 1 MSEKNFVITPIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVFLYLTGDEHGQKIQK 60
Db 1 MAKETFTVITPIYPSGNLHIGAYSTVAGDVLSRYKMGQYDVFLYLTGDEHGQKIQK 60

Qy 61 ABEAGITPQAYVDMGVKELQQLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYL 120
Db 61 AQXAGKTELEYLDEMISGLSKLSKLEISNDDFIRTTDRHKQVKEVFERLLKQGGIYL 120

Qy 121 GEYSGWYSVDEBFFTESQLAEVFRDEAGNVGTGIAPSGHEVWSEESYFLRLSKYQD 179
Db 121 GEYEGWYSVDEYTYTYSQVDPYI-ENKIVGKSPDSGHEVELVKEESYFFNINKYTD 179

Qy 180 RLVEFFKAHPFETPDGRLNEMLRNFIPLGLEDLAVSRTTFTWGPVPSNPKHVVYWD 239
Db 180 RLLEFFDANPDFIOPPSRKNEMINNFIPGLLEDLAVSRTSFDWGVVPSNPKHVVYWD 239

Qy 240 ALLNATAGYAOEHGDFKFWGTWFMVYKGLDPLRYLMRLNLPVSGDGTFTPEDYVGRIN 299
Db 240 ALVNYISALGYLSDDELQKYPADV-HLMAKEIVRFHSIIWPIILLMALDLPKPKVFA 298

Qy 300 HGMVFMKDGKMSKSGNVVPEMLVRYGLDPLRYLMRLNLPVSGDGTFTPEDYVGRIN 359
Db 299 HGMILMKDGKMSKSGNVVDPVLIIDRYGLDATRYVLMELPFGSDGVFTPEAFVETNY 358

Qy 360 ELANDLGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADFTHTMEAVDIP 419
Db 360 ELANDLGNLNRVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEKSTADFTHTMEAVDIP 419
```

Db 359 DLANDLGNLVRTISNINKYFOGELPAYEGPKHELDMEALAHETVXHFNMESFQFS 418
 QY 420 RALEAVWTLISRTNKYIDTAPWLDKDBALDQLASVNSHWQASIRVVAHLIEPMMET 479
 Db 419 VALSTVWKFEISRTNKYIDTETPEWVLAKDSQKMDLGNVMAHLVENIRFAAVLLRPFLTHA 478
 QY 480 SRANVLTOIGLEVSLENISLADFP--DVTVAAGTPIFPRLANEHEIAYIKEQMBGN 536
 Db 479 PKEIFQLNINPELPELESQYALKQPIWVTEKPTPIFPRLTEAEIAYIKESMOPP 538
 QY 537 KPAVEKEMPDEVELKLNKDEIKFDFKVEIRVAEVKEVSKEVSGDKLLQFRLDAGDGE 596
 Db 539 K----SEESKDEVE-RPSKAQIDIKDFDKVEIKAAITIDAENVKPSDKLLKIQIDLGL-LE 592
 QY 597 DROILSGIAKYPNEBELVKKVQIVANLKPRKMKKYVSQGMILSAEHDGKLTLLTVDP 656
 Db 593 QRCIVSGIAKFY-RPEDIIGKVVAVTNLKPAPKLMQ-KSEGMILSAEKGDTLLVSUPS 650
 QY 657 AVPNGSVI 664
 Db 651 AIPNGAVI 658

RESULT 12
 US-10-282-122A-46123
 ; Sequence 46123, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
 CURRENT FILING DATE: 2003-02-20
 PRIOR FILING DATE: 2000-03-21
 PRIOR FILING DATE: 2000-05-23
 PRIOR FILING DATE: 2000-05-26
 PRIOR FILING DATE: 2000-09-06
 PRIOR FILING DATE: 2000-09-09
 PRIOR FILING DATE: 2000-10-23
 PRIOR FILING DATE: 2000-11-27
 PRIOR FILING DATE: 2000-12-22
 PRIOR FILING DATE: 2001-02-09
 PRIOR FILING DATE: 2001-02-16
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 78614
 -SOFTWARE: Patent in version 3.1
 SEQ ID NO 46123
 LENGTH: 669
 TYPE: PRT
 ORGANISM: Bacillus anthracis
 US-10-282-122A-46123

Query Match 55.6%; Score 1935.5; DB 15; Length 669;
 Best Local Similarity 57.0%; Pred. No. 2.7e-142;
 Matches 383; Conservative 107; Mismatches 161; Indels 21; Gaps 12;
 QY 1 MSEK--FYITPTIYPSGKLHIGSATTTACDVLAARYKELMGYDVLYLGLDEHGOKIQ 58
 Db 10 MTEENKSFYITPTIYPSGKLHIGHAYTTVAGDAMARYKCMQGYNVHYLTGTDEHGOKIQ 69
 QY 59 QKAEAGITTPQAYVDGMVGVKELWQLLDISYDKFIRTTDDYHKKVVAQVFERLLAQDDI 118
 Db 70 KKAELNVTPOAYVDNIVAGIKELWEXMDISYDDFTTDRHKVDVEKIFKQLVDQGD 129
 QY 119 YLGEYSGMYSDEEFFTESQLAEVFRDEAGNVTGGIAP--SCHEVEVWVSSESTFLRLSKY 177
 Db 130 YLDEVEGWYSVQDETFTYVTEHQLVDPIM--EGDKVVGKSPSDGHDVLRRESEYFFRMKY 188
 QY 178 QDLVEFEKAPBEITPDGRNLNEMRNFIIPGLEDLAVSRRTFTWGVVPSNPKHVVYVW 237
 Db 189 VDRLLKPYEDNPHFIQEPESRKNEMINNFIRPGLLEDLAVSRITSFDGKGVVRVPGNPKHVIYVW 248
 QY 238 IDALLNATAGLYAQDBEHNFGTVPFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRL 297
 Db 249 VDALSNTYITAGYGTENEEMYKFWPADV--HLVCKEIVRPHTIYWPILLMALDLPKPKV 307
 QY 298 IAHGFWNKGKMSKSGNVTYVPEMLVERIGLDPIRYILMNLVPGSDGTFTPDYVGR 357
 Db 308 FAHGWLKMDGMSKSGNVVDPVILIDRYGLDALRYLLREVDFGSDGVPTPEGFVER 367
 QY 358 NYELANDLGNLNLRTVSMINKYFDGOIPAYVEGVTPEFDHVLAEVAEKSIADPFHTMEAVD 417
 Db 368 NFDLANDLGNLNLRTVAMIDKYFNGEIPAFKANVTEDETLVTFKDTLKKVEEAMENNE 427
 QY 418 YPRALEAVWTLISRTNKYIDTAPWLDKDBALDQLASVNSHWQASIRVVAHLIEPMM 477
 Db 428 FVALSSTIQLVSRNKYIDETQPVWLAKDEDDREKLASVMAHLAEVLRTGTGIMLMPFLT 487
 QY 478 ETSRAVLTOIGL--REVSSLENLS--LADFPADVTVAAGTPIFPRLNMEESIAYIKEQME 534
 Db 488 VAPSKQFAQLGLTDEAHTSWGSLSTIGICIPAG--TKVEKGQPIFRLNDVVAIKEQMK 546
 QY 535 GNKPAVE--KEWNPDEVELKLNKDBEIKFDFKVEIRVAEVKEVSKEVSGDKLLQFLRDA 592
 Db 547 ASAPKVEEKKEEP-----KAEITIDDFKVELRVAEVLAEVLSAEPVKADKLLKIQIDL 599
 QY 593 GDGEDROILSGIAKYPNEBELVKKVQIVANLKPRKMKKYVSQGMILSAEHDGKLTLL 652
 Db 600 GT-EXRQVVGIAKFY-SPEDLKGGKKVICVNLKPVK--LRGELSQGMILAGEENGVLSDA 656
 QY 653 TVDPAVPNGSVI 664
 Db 657 SIDQNPNGTKI 668

RESULT 13
 US-09-815-242-12135
 ; Sequence 12135, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR FILING DATE: 2000-03-21

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12135
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12135

Query Match          55.6%; Score 1934; DB 9; Length 657;
Best Local Similarity 56.3%; Pred. No. 3.4e-142;
Matches 376; Conservative 111; Mismatches 165; Indels 16; Gaps 10;

QY      1 MGEKNFYITPIYPGSKLHIGSAVTIACDVLARVYKRLMGVDVFLYTLGLDEHGQKIQOK 60
Db      1 MAKETEYITPIYPGSKLHIGSAVTIACDVLARVYKRLMGVDVFLYTLGLDEHGQKIQOK 60

QY      61 ABEAGITPQAYVDGMVAGVKELWQLLIDISYDKPIRTDDYHKVVAQVPERLLAODDIYL 120
Db      61 ABEAGITPQAYVDGMVAGVKELWQLLIDISYDKPIRTDDYHKVVAQVPERLLAODDIYL 120

QY      61 AOKAGKTEYLDDEMIAGIKQLWAKLEISNDPFIKRLMGVDVFLYTLGLDEHGQKIQOK 120
Db      61 AOKAGKTEYLDDEMIAGIKQLWAKLEISNDPFIKRLMGVDVFLYTLGLDEHGQKIQOK 120

QY      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179
Db      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179

QY      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179
Db      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179

QY      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239
Db      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239

QY      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239
Db      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239

QY      240 ALLNYATAGYAQDEHGNFDFKNGVTFHMGVKDILRFHSIYWPILLMMLDVKLDPRLIA 299
Db      240 ALLNYATAGYAQDEHGNFDFKNGVTFHMGVKDILRFHSIYWPILLMMLDVKLDPRLIA 299

QY      240 ALVNYISALGYLSDDSESLFNKYWPADI-HLMAKEIVRFHSIYWPILLMMLDVKLDPRLIA 298
Db      240 ALVNYISALGYLSDDSESLFNKYWPADI-HLMAKEIVRFHSIYWPILLMMLDVKLDPRLIA 298

QY      300 HGVFVMDGKMSKSGNVVYPMELVRYGLDPLRYLLMRNLPVGSDDGTFTPEDYVGRIN 359
Db      300 HGVFVMDGKMSKSGNVVYPMELVRYGLDPLRYLLMRNLPVGSDDGTFTPEDYVGRIN 359

QY      299 HGVFVMDGKMSKSGNVVYPMELVRYGLDPLRYLLMRNLPVGSDDGTFTPEDYVGRIN 358
Db      299 HGVFVMDGKMSKSGNVVYPMELVRYGLDPLRYLLMRNLPVGSDDGTFTPEDYVGRIN 358

QY      360 ELANDLGNLNTVSMINKYFGQIPAYVEGTEDPHVLAEVAEKSIADFTHEANDYP 419
Db      360 ELANDLGNLNTVSMINKYFGQIPAYVEGTEDPHVLAEVAEKSIADFTHEANDYP 419

QY      359 DLANDLGNLNTVSMINKYFGQIPAYVEGTEDPHVLAEVAEKSIADFTHEANDYP 418
Db      359 DLANDLGNLNTVSMINKYFGQIPAYVEGTEDPHVLAEVAEKSIADFTHEANDYP 418

QY      420 RALEAVWTLSIRTKYIDETAPWLKDDEALDQLASVMSHWQASIRVVVAHLIRPFVME 479
Db      420 RALEAVWTLSIRTKYIDETAPWLKDDEALDQLASVMSHWQASIRVVVAHLIRPFVME 479

QY      419 VALSITWKEISIRTKYIDETAPWLKDDEALDQLASVMSHWQASIRVVVAHLIRPFVME 478
Db      419 VALSITWKEISIRTKYIDETAPWLKDDEALDQLASVMSHWQASIRVVVAHLIRPFVME 478

QY      480 SRAVITQGLEEVSSLENLSADP---PADVTVVAKGTPPIPRLNMBEIIAYIKQEMGN 536
Db      480 SRAVITQGLEEVSSLENLSADP---PADVTVVAKGTPPIPRLNMBEIIAYIKQEMGN 536

QY      479 PKEIFEQLINNPFQMEFSSLAQYGVLTPEIPIVTPQPKPIPRLDSEAEIAYIKESMQ-- 536
Db      479 PKEIFEQLINNPFQMEFSSLAQYGVLTPEIPIVTPQPKPIPRLDSEAEIAYIKESMQ-- 536

QY      537 KPAVKEMNPDVELKLNKDEIKFDEDEKVEIRVAEVEKSVESGDKLLOFRLDAGDGE 596
Db      537 KPAVKEMNPDVELKLNKDEIKFDEDEKVEIRVAEVEKSVESGDKLLOFRLDAGDGE 596

QY      537 PRATEE--KEIP---SKPDIKDFKVELIKATIIDAHEVKSCKLLKIQVLDL-DSE 590
Db      537 PRATEE--KEIP---SKPDIKDFKVELIKATIIDAHEVKSCKLLKIQVLDL-DSE 590

QY      597 DRQILSGIAKYPNQELVGKVKQIVANLKPRLMKKTVSQMILSABHGKLLITVDP 656
Db      597 DRQILSGIAKYPNQELVGKVKQIVANLKPRLMKKTVSQMILSABHGKLLITVDP 656

QY      591 QRQIVSGIAKYP--TPDDIIGKKVAVVTLNKPAPLMGQ--KSEGMILSABKGVLVLSLPS 648
Db      591 QRQIVSGIAKYP--TPDDIIGKKVAVVTLNKPAPLMGQ--KSEGMILSABKGVLVLSLPS 648

QY      657 AVPNGSVI 664
Db      657 AVPNGSVI 664

QY      649 ALPNGAVI 656
Db      649 ALPNGAVI 656

RESULT 14
US-10-282-122A-43933
; Sequence 43933, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 43933
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43933

Query Match          55.5%; Score 1931; DB 15; Length 657;
Best Local Similarity 56.3%; Pred. No. 5.8e-142;
Matches 377; Conservative 110; Mismatches 163; Indels 20; Gaps 11;

QY      1 MGEKNFYITPIYPGSKLHIGSAVTIACDVLARVYKRLMGVDVFLYTLGLDEHGQKIQOK 60
Db      1 MAKETEYITPIYPGSKLHIGSAVTIACDVLARVYKRLMGVDVFLYTLGLDEHGQKIQOK 60

QY      61 ABEAGITPQAYVDGMVAGVKELWQLLIDISYDKPIRTDDYHKVVAQVPERLLAODDIYL 120
Db      61 ABEAGITPQAYVDGMVAGVKELWQLLIDISYDKPIRTDDYHKVVAQVPERLLAODDIYL 120

QY      61 AOKAGKTEYLDDEMIAGIKQLWAKLEISNDPFIKRLMGVDVFLYTLGLDEHGQKIQOK 120
Db      61 AOKAGKTEYLDDEMIAGIKQLWAKLEISNDPFIKRLMGVDVFLYTLGLDEHGQKIQOK 120

QY      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179
Db      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179

QY      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179
Db      121 GEYSGWYSVDEFFTESQLAEVFRDEAGNVGGIAP-SGHEVEWVSESYFLRLSKYQD 179

QY      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239
Db      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239

QY      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239
Db      180 RLVFFPKAHPFITPDGRINEMLRNFIKPLGLADLAVSRTEFWGVPSNPKHVVYVWID 239

QY      240 ALLNYATAGYAQDEHGNFDFKNGVTFHMGVKDILRFHSIYWPILLMMLDVKLDPRLIA 299
Db      240 ALLNYATAGYAQDEHGNFDFKNGVTFHMGVKDILRFHSIYWPILLMMLDVKLDPRLIA 299

QY      240 ALVNYISALGYLSDDSESLFNKYWPADI-HLMAKEIVRFHSIYWPILLMMLDVKLDPRLIA 298
Db      240 ALVNYISALGYLSDDSESLFNKYWPADI-HLMAKEIVRFHSIYWPILLMMLDVKLDPRLIA 298
```

QY 300 HGWFVKDGKSKGNVYVPEMLVERGLDPLRYLLMRNLPGSGCTFTPDYGRINY 359
DB 299 HGWILKDGKSKGNVVDNLIIDRYGLDTRYLMLRELPGSGDGVFTPEAFVERTNF 358
QY 360 ELANDLGNLLNRTVSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADFHTHMEAVDYP 419
DB 359 DLANDLGNLVNRTISWNKYFDGELPAYQGPLHELDDEMEVALETVKSYTESMESLQFS 418
QY 420 RALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPMMET 479
DB 419 VALSTVWKTFISRTNKYIDETTPWVLAKDQSKDMLGNVMAHLVENIRYAVALRPLFLTHA 478
QY 480 SRAVLTOIGLE-----EVSSLENLSLADFPADVTVVAKGTPIPPRLNMEEBIAYIKEOME 534
DB 479 PKEIFEQLNINNPQFMESLSQYGVN--LTSIMVTGQPKPIPPRLDSAEIAYIKESNQ 536
QY 535 GNKPAVEKWNPDVELKLNKDIKPEDFKVEIRVAEVEKSVKSGDKLLQFRLDAGD 594
DB 537 --PPATEEE--KEEIP---SKPQIDIKDFDKVEIKAATIIDAEHVKSKDKLLKIQVDL-D 588
QY 595 GEDROIILSGIAKYYPNEQELVGKKVQIVANLKPROMKKYVSGMILSAEHDGKLTLLTV 654
DB 479 PKEIFEQLNINNPQFMESLSQYGVN--LTSIMVTGQPKPIPPRLDSAEIAYIKESNQ 536
QY 589 SEQROIIVSGIAKFY--TPDDIIGKKVAVVTNLKPAKLMGQ--KSEGMILSAEKGVLTLVSL 646
DB 655 DPAVNGSVI 664
DB 647 PSAIPNGAVI 656

RESULT 15

US-09-815-242-13013
; Sequence 13013, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/131,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13013
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13013

Query Match 55.5%; Score 1930; DB 9; Length 657;
Best Local Similarity 56.1%; Pred. No. 6,9e-142;
Matches 376; Conservative 108; Mismatches 166; Indels 20; Gaps 9;

QY 1 MSEKNFVITTPYPPSGKLHIGSAYTTIACDLVARYKRLMGYDVYLTGLDEHGKIQOK 60
DB 1 MAKTFITTPYPPSGNLHIGHAYSVAGDVARYKEMQGYDVRILTGTDEHGKIQEK 60
QY 61 AEEAGITPOQAYVDMAYGVXELMQLLDISYDKFTRTDDYHEKVAQVFFRLLAODIYL 120
DB 61 AOKAGKTEIEYLDDEMIAGIKQWAKLEISNDFTIRTEERHKHYVEQVFFRLLAODIYL 120
QY 121 GEYSGWYSVDEBEPFTESQLAEVFRDEAGNVGTGIAP--SGHEVEWVSEESYFFRLSKYQD 179
DB 121 GEYEGWYSVPDETYYTSSQLVDP--OYENGKLIIGKSPDSGHEVELVKEESYFFNISKYTD 179
QY 180 RLVEFFFKAHPEFTITPDGRLEMLENFIEPGLLEDLAVSRTTFTWGVPPVPSNPKHVVYVWD 239
DB 180 RLLEFYDQNPDIQPPSKQENMINNFIKPGLADLAVSRTSFNMGVHVPSPKHVVYVWD 239
QY 240 ALLNYATRALGVAQDEHGNFDFKFWNGTVFHMVKDILBFHSIYWPILLMLLDVLPDLRLIA 299
DB 240 ALVNYISALGYSDDDESLEFNKYWPADI--HLMAKSIVFRHSIIFWILLMALDLPKPKVFA 298
QY 300 HGWFVKDGKSKGNVYVPEMLVERGLDPLRYLLMRNLPGSGDGTFTPDYGRINY 359
DB 299 HGWILKDGKSKGNVVDNLIIDRYGLDTRYLMLRELPGSGDGVFTPEAFVERTNF 358
QY 360 ELANDLGNLLNRTVSMINKYFDGOIPAYVEGVTEFDHVLAEVAEKSIADFHTHMEAVDYP 419
DB 359 DLANDLGNLVNRTISWNKYFDGELPAYQGPLHELDDEMEVALETVKSYTESMESLQFS 418
QY 420 RALEAVWTLISRTNKYIDETAPWVLKDDEALRDQLASVMSHWQASIRVVAHLIEPMMET 479
DB 419 VALSTVWKTFISRTNKYIDETTPWVLAKDQSKDMLGNVMAHLVENIRYAVALRPLFLTHA 478
QY 480 SRAVLTOIGLE-----EVSSLENLSLADFPADVTVVAKGTPIPPRLNMEEBIAYIKEOME 534
DB 479 PKEIFEQLNINNPQFMESLSQYGVN--LTSIMVTGQPKPIPPRLDSAEIAYIKESNQ 536
QY 535 GNKPAVEKWNPDVELKLNKDIKPEDFKVEIRVAEVEKSVKSGDKLLQFRLDAGD 594
DB 537 PPATKEEKEEIP-----SKPQIDIKDFDKVEIKAATIIDAEHVKSKDKLLKIQVDL-D 588
QY 595 GEDROIILSGIAKYYPNEQELVGKKVQIVANLKPROMKKYVSGMILSAEHDGKLTLLTV 654
DB 589 SEQROIIVSGIAKFY--TPDDIIGKKVAVVTNLKPAKLMGQ--KSEGMILSAEKGVLTLVSL 646
QY 655 DPAVNGSVI 664
DB 647 PSAIPNGAVI 656

Search completed: December 17, 2004, 22:11:32
Job time : 151 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:51:08 ; Search time 42 Seconds
(without alignments)
1523.432 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITTPYPSGKLH.....DGKLTLLITVDPAVNGSVIG 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3446	99.1	665	2 D95091	methionyl-tRNA syn
2	3446	99.1	679	2 H97958	methionine-tRNA 11
3	2392.5	68.8	662	2 H86723	methionine-tRNA 11
4	2113	60.7	664	2 AB1097	methionyl-tRNA syn
5	2107	60.6	664	2 A11459	methionyl-tRNA syn
6	1983.5	56.4	664	2 S66067	methionine-tRNA 11
7	1983.5	56.2	660	2 E83856	methionyl-tRNA syn
8	1931	55.5	657	2 C89815	methionyl-tRNA syn
9	1919	55.2	649	2 S16682	methionine-tRNA 11
10	1640.5	47.2	644	2 A97468	methionyl-tRNA syn
11	1359.5	39.1	629	2 E72297	methionine-tRNA 11
12	1189.5	34.2	628	2 F81356	methionyl-tRNA syn
13	1183.5	33.2	621	2 C82922	methionyl-tRNA syn
14	1152.5	33.1	656	2 H71867	methionine-tRNA 11
15	1148.5	33.0	650	2 A64572	methionine-tRNA 11
16	1140	32.8	497	2 G70408	methionine-tRNA 11
17	1131	32.5	508	2 F97830	methionine-tRNA 11
18	1129	32.5	616	1 SVTWT	methionyl-tRNA syn
19	1068.5	30.7	516	2 P97541	methionyl-tRNA syn
20	1068.5	30.7	516	2 A82760	methionine-tRNA 11
21	1062	30.5	508	2 F71674	methionine-tRNA 11
22	1058	30.4	544	2 AB3375	methionyl-tRNA syn
23	1047.5	30.1	530	2 A11835	probable methionyl
24	1040	29.9	537	2 F86938	methionine-tRNA 11
25	1037.5	29.8	532	2 S76525	probable methionyl
26	1037	29.8	519	2 B70603	methionyl-tRNA syn
27	1018	29.3	681	2 C75395	hypothetical prote
28	974	28.0	509	2 B90573	methionine-tRNA 11
29	904	26.0	616	2 T47679	methionine-tRNA 11

RESULT 1

D95091
methionyl-tRNA synthetase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: D95091
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.
nson, T.; Hickey, B.K.; Holt, I.E.
Science 293 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: D95091
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-665 <KJF>
A:Cross-references: UNIPROT:Q97RL6; GB:AB005672; PIDN:AAK74925.1; PID:g14972263; GSPDB:G
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0788
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 99.1%; Score 3446; DB 2; Length 665;
Best Local Similarity 99.2%; Pred. No. 2.3e-198;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY	1	MSEKNFYITTPYPSGKLHIGSAVTTIACDVLARYKELMGYDVFLTGLDEHGKIQOK	60
DB	1	MSEKNFYITTPYPSGKLHIGSAVTTIACDVLARYKELMGYDVFLTGLDEHGKIQOK	60
QY	61	ABEAGITQAYVDGMVAVGKELWQLLDISYDKFRTTDDYHEKVAQVFERLLAODDIYL	120
DB	61	ABEAGITQAYVDGMVAVGKELWQLLDISYDKFRTTDDYHEKVAQVFERLLAODDIYL	120
QY	121	GEYSGWYSVDSDEFTESQAEVFRDEAGNVTGGIAPSGHEVEMWSEESYFLRLSKYQDR	180
DB	121	GEYSGWYSVDSDEFTESQAEVFRDEAGNVTGGIAPSGHEVEMWSEESYFLRLSKYQDR	180
QY	181	LVEFFKAHPEFTTPGRLNEMLRNFIKGLDLAVSRRTFTWGVPSNPKHVVVWIDA	240
DB	181	LVEFFKAHPEFTTPGRLNEMLRNFIKGLDLAVSRRTFTWGVPSNPKHVVVWIDA	240
QY	241	LNNYATLGYAODEHGNFDFKFWNGVFMVGVKDIIRFHSIYWPILLMMLDVKLPLRLAH	300
DB	241	LNNYATLGYAODEHGNFDFKFWNGVFMVGVKDIIRFHSIYWPILLMMLDVKLPLRLAH	300
QY	301	GWFMKDGKMSKSGNYYVPEMLVRYGLDPLRYLYLMENLPVSGDTTPEDYVGRINYE	360
DB	301	GWFMKDGKMSKSGNYYVPEMLVRYGLDPLRYLYLMENLPVSGDTTPEDYVGRINYE	360
QY	361	LANDIGNLLNRVSMINKYFDQIQIYAYVGVTEFDFHVLAEVAKSIADPHTHWEAVDYPR	420

```
Db 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTHMEADVYPR 420
QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
Db 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
QY 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEYKESKVEGSDKLLQFRLDAGDGEDRQI 600
Db 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEYKESKVEGSDKLLQFRLDAGDGEDRQI 600
QY 601 LSGIAKYYPNEQELVGKQVQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 601 LSGIAKYYPNEQELVGKQVQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
QY 661 GSVIG 665
Db 661 GSVIG 665

RESULT 2
H97958
methionine-tRNA ligase (EC 6.1.1.10) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 23-Dec-2002
C:Accession: H97958
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burchett, S.; DeHoff, B.S.; H
y, P.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.; M
J. Bacteriol. 183, 5703-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: H97958
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-679 <KOR>
A:Cross-references: GB:AE007317; PIDN:AAK99500.1; PID:g15458285; GSPDB:GN00174
C:Genetics:
C:Gene: metS
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: ligase

Query Match 99.1%; Score 3446; DB 2; Length 679;
Best Local Similarity 99.2%; Pred. No. 2.4e-198;
Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 60
Db 15 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 74
QY 61 ABEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAODDIYL 120
Db 75 ABEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAODDIYL 134
QY 121 GYSYSGWYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
Db 135 GYSYSGWYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 194
QY 181 LVEFFKAHPEFTTPDGRNLEMLRNIEPGLDLAVSRITFTTGVGPVPSNPKHVYVWIDA 240
Db 195 LVEFFKAHPEFTTPDGRNLEMLRNIEPGLDLAVSRITFTTGVGPVPSNPKHVYVWIDA 254
QY 241 LLNATYALGYAODEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 300
Db 255 LLNATYALGYAODEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 314
QY 301 GWFVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTTPDYVGRINYE 360
```

```
Db 315 GWFVMDGKMSKGNVYVPEMLVERYGLDPLRYLLMRNLPGVSDGTFTTPDYVGRINYE 374
QY 361 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTHMEADVYPR 420
Db 375 LANDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAEQSIADPHTHMEADVYPR 434
QY 421 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 480
Db 435 ALBAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMSHMOASIRVVAHLIEPFMMETS 494
QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 540
Db 495 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIFPRLNMEEEIAYIKEQMEGNKPAV 554
QY 541 EKWNPDVELKLNKDEIKFEDFKVEIRVAEYKESKVEGSDKLLQFRLDAGDGEDRQI 600
Db 555 EKWNPDVELKLNKDEIKFEDFKVEIRVAEYKESKVEGSDKLLQFRLDAGDGEDRQI 614
QY 601 LSGIAKYYPNEQELVGKQVQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 660
Db 615 LSGIAKYYPNEQELVGKQVQIVANLKPRKMKKKYVSQGMILSAEHDGKLTLLTVDPAPVN 674
QY 661 GSVIG 665
Db 675 GSVIG 679

RESULT 3
H86723
methionine-tRNA ligase (EC 6.1.1.10) [imported] - Lactococcus lactis subsp. lactis (strai
N:Alternate names: methionyl-tRNA synthetase
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86723
R:Boloian, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: H86723
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-662 <STO>
A:Cross-references: UNIPROT:Q9CHE0; GB:AE005176; PID:g12723712; PIDN:AAK04890.1; GSPDB:GN
A:Experimental source: strain IL1403
C:Genetics:
C:Gene: metS
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: ligase

Query Match 68.8%; Score 2392.5; DB 2; Length 662;
Best Local Similarity 68.6%; Pred. No. 2.3e-135;
Matches 459; Conservative 88; Mismatches 113; Indels 9; Gaps 7;
QY 1 MSEKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 59
Db 1 MTEKTFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 60
QY 60 KAEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAODDIY 119
Db 61 KAEAGITPOAYVDGNAVGKELWQLLDISYDKFIRTTDDYHEKVVAQVFERLLAODDIY 120
QY 120 GYSYSGWYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 179
Db 121 LGKYAGWYSVDSDEFFTESQAEVFEDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDR 180
QY 180 RLVSEFFKAHPEFTTPDGRNLEMLRNIEPGLDLAVSRITFTTGVGPVPSNPKHVYVWIDA 239
Db 181 WLTQYIDEPDFIOPVPRKMNWNNIFKGLDLAIVSRITFTTGVGPVPSNPKHVYVWIDA 240
QY 240 ALLNATYALGYAODEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 299
Db 241 ALLNATYALGYAODEHGNFDFKFWNGTVFHMVGKDIILRFHSIYWPILLMMLDKVLPDLIAH 298
```

QY 300 HGFWVKDGMKSGKGVVYPEMLVERYGIDPLRYVYLMRNLVPGVSGTTPEDYVGRINY 359
DB 299 HGWLLKDGKMSKGVVYPEMLIBRYGLDAVRYVYLMRAISFGQDGIPTDFVGRINF 358
QY 360 ELANDLGNLNRVSMINKYDFGQIIPAYVEGV--TEPDHVLAEVAEKSIADPFTHEAVDY 418
DB 359 DLANDLGNLNRVSMINKYNDGKIEA--TGVSTEDASLEEVETISHFHKAMDKPEF 416
QY 419 PRALAVWTLISRTNKYIDETAPWVLDKDBALRDLQASVSHWQASIRVVVAHLIPEPFME 478
DB 417 NVRALADVWTLISRTNKYIDETAPWVLAKEBDDKAKLNNVLYHLAENLRAGALLQPFMRA 476
QY 479 TSRAVLTQIGLEBVS--SLENLSLADFPADVTVVAKGTPFPRLNMBEEIAYIKQOMEGNK 537
DB 477 TSKIEFQGLNDRSRLSLENLSFG--YSFTHPVVAKGQTFPRLDVBEVYAYIKLQWAGG- 534
QY 538 PAYEKWNPDEVELKLNKDEIKFEDPKVEIRVAEYKESKVEGSKLQFLRDAGDGB 597
DB 535 VLPEKWPPEVELNLTLPQIKEDDEKTELKVAEVEPEVSGDKLRFKLDAGDSBP 594
QY 598 RQILSGIAKYRNEOELVGVKQVIVANLKPRKMKKYVSGMILSAFHGKTLITVDPA 657
DB 595 RQILSGIAQYPNEQELVGVKQVIVANLKPRKMKKYVSGMILSAEFGKLSVLTVD 654
QY 658 VPNGSVIG 665
DB 655 VPAGSLIG 662

RESULT 4
AB1097
methionyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1097
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1097
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <GLA>
A:Cross-references: UNIPROT:Q9APZ; GB:NC_003210; PID:GAC98392.1; PID:G16409534; GSPDB:
A:Experimental source: strain EGD-e
C:Gene: mets
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 60.7%; Score 2113; DB 2; Length 664;
Best Local Similarity 60.7%; Pred. No. 1.2e-118;
Matches 407; Conservative 101; Mismatches 144; Indels 18; Gaps 8;

QY 242 LNYATAGYAQDEHNGNFKFNGTVPFHWGKDIILRHSIYWPILLMMLDVKLDPRLIAHG 301
DB 244 SNYITAGYNTDNTKFKYMPADV-QIVGKEIVRFHTIYWPIMLMALDPLPKMVFHG 302
QY 302 WFWKDGKMSKGVVYPEMLVERYGIDPLRYVYLMRNLVPGVSGTTPEDYVGRINYEL 361
DB 303 WILKDGKMSKGVVDPYMLIDRYGLDALRYVLLREVPGSGDLFTPEFVDRVNF 362
QY 362 ANDLGNLNRVSMINKYDFGQIIPAYVEGVTEPDHVLAEVAEKSIADPFTHEAVDY 421
DB 363 ANDLGNLNRVSMINKYDFGQIIPAYQNVTEPFQIILVDFKNNVYKVEYEGMDHMFQSV 422
QY 422 LEAVWTLISRTNKYIDETAPWVLDKDBALRDLQASVSHWQASIRVVVAHLIPEPFMETS 481
DB 423 LNQWLSLISRTNKYIDETAPWALAKEDKRTSLASVTHLAENLRIIAVLLQFLTP 482
QY 482 AVLTIQIGLEEVSSLENL-----SLADFPADVTVVAKGTPFPRLNMBEEIAYIKQOM 534
DB 483 EIFLQGLQE---ENLKKWDSIYGYEIPAGTIVVKKGTPIFPRDLAEVETVIQDEMK 538
QY 535 GNKPAVEKWNPDVELKLNKDEIKFEDPKVEIRVAEYKESKVEGSKLQFLRDAGD 594
DB 539 GSAPAPABEE--VAEVE-ALETPQIGIEDFKIDLRVAEVKQDVKKKADKLLCFOLD 595
QY 595 GEDRQILSGIAKYRNEOELVGVKQVIVANLKPRKMKKYVSGMILSAEHGKTLITV 654
DB 596 GKLRQVLSGIAEFYQPE-ELIGKKVIVVSNLKPVK-LRGLMSEGMILSGEKDKLSVIE 653
QY 655 DPVNGSVI 664
DB 654 SSALPNGAKV 663

RESULT 5
AB1459
methionyl-tRNA synthetase [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1459
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.; Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1459
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <GLA>
A:Cross-references: UNIPROT:Q92P90; GB:AL592022; PID:GAC95449.1; PID:G16412635; GSPDB:G
A:Experimental source: strain Clip11262
C:Gene: mets
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 60.6%; Score 2107; DB 2; Length 664;
Best Local Similarity 60.6%; Pred. No. 2.7e-118;
Matches 406; Conservative 101; Mismatches 145; Indels 18; Gaps 8;

QY 3 EKN-FYITPTIYPSGKLHIGSAVTTACDVLARYKRLMGYDVLYLTGLDEHGQKIQA 61
DB 4 EKNTFYITPTIYPSGKAHIGHAYTTVAGDAMARYKRLKGYDVLYLTGTDHGQKIQA 63
QY 62 BEAGITPOAYVDGMAGVKELWQLLDISYDXFIRTTDDYHEKVAQVFERLLAQDDIYLG 121
DB 64 KERGISEQYVDIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEGGDIYLG 123
QY 122 EYSGWYSVDSDEEFTTESQLAEVFRDEAGNVTGGTAPSGHEVWVSEBSYFRLSKYQDRL 181
DB 124 EYEGWYSVDSDEEFTTESQLAEVFRDEAGNVTGGTAPSGHEVWVSEBSYFRLSKYQDRL 183
QY 182 VEPFKAHPEITPDGRINLMLNFIPLGLEDLAVSRRTFTWGVVPSNPKHVVYVWIDAL 241
DB 184 VEYNGHPEPILPESRKNNMINNFIPKPLEDLAVSRRTFTWGVVPSNPKHVVYVWIDAL 243

QY 3 EKN-FYITPTIYPSGKLHIGSAVTTACDVLARYKRLMGYDVLYLTGLDEHGQKIQA 61
DB 4 EKNTFYITPTIYPSGKAHIGHAYTTVAGDAMARYKRLKGYDVLYLTGTDHGQKIQA 63
QY 62 BEAGITPOAYVDGMAGVKELWQLLDISYDXFIRTTDDYHEKVAQVFERLLAQDDIYLG 121
DB 64 KERGISEQYVDIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEGGDIYLG 123
QY 122 EYSGWYSVDSDEEFTTESQLAEVFRDEAGNVTGGTAPSGHEVWVSEBSYFRLSKYQDRL 181
DB 124 EYEGWYSVDSDEEFTTESQLAEVFRDEAGNVTGGTAPSGHEVWVSEBSYFRLSKYQDRL 183

[illegible]

RESULT 6

S66067 methionine-tRNA ligase (EC 6.1.1.10) mets - Bacillus subtilis
N;Alternate names: methionyl-tRNA synthetase mets
C;Species: Bacillus subtilis
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66067; E69657
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chromo-
A;Reference number: S65967; MUID:96051385; PMID:7584024
A;Accession: S66067
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-664 <OGA>
A;Cross-references: UNIPROT:P37465; EMBL:D26185; NID:9467326; PIDN:BAAC05273.1; PID:94674
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ched-
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Faret, C.; Ferrari, E.;
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler-
iech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.P.;
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel-
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, I.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Segiguchi, J.; Sekowska, A.; Seron-
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69657
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

A;Residues: 1-664 <KUN>
A;Cross-references: GS:Z99104; GB:AL009126; NID:G2632267; PIDN:CAB11814.1; PID:G2632305
A;Experimental source: strain 168
C;Genetics:
A;Gene: mets
C;Superfamily: methionyl-tRNA synthetase, dimer-forming
C;Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

RESULT 7
E83656
methionyl-tRNA synthetase mets [imported] - Bacillus halodurans (strain C-125)
C#Species: Bacillus halodurans
C#Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C#Accession: E83656
P;Takai, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirano, T.
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; M01D:20512582; PMID:11058132
A;Accession: E83656
A;Status: preliminary

A:Accession: C98915
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-657 <KUR>
A:Cross-references: UNIPROT:Q99WB3; GB:BA000018; PID:gi3700380; PIDN:BAB41678.1; GSPDB:G8
A:Experimental source: strain N315
C:Genetics:
A:Gene: metS
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 55.5%; Score 1931; DB 2; Length 657;
Best Local Similarity 56.3%; Pred. No. 8.e-108;
Matches 377; Conservative 110; Mismatches 163; Indels 20; Gaps 11;

QY 1 MSEKTYIITPIYPYSGKLHGSAVTTACDVLARYKRLMGDYDFYLTCGLDEHGQKIQQK 60
DB 1 MAKETVIITPIYPYSGNLHGHAVSTVAGDVIAIKRMQGYDVRILGTDEHGQKIQEK 60

QY 61 AEEAGITPOAYVDGMAGVKELWQLLDTSYDKFIPTDDYHEKVAQVFERRLLAODDIYL 120
DB 61 AQAKAGTEIYLEDMLAGIKQWLAKLEISNDPFIETTERHKHVVEQVFERLLKGGDIYL 120

QY 121 GEYSGWYSDEBEFFTESQLAEVFDEAGNVTGGIAP--SGHEVEMVSBESEYFLRLSKTQD 179
DB 121 GEYEGWYSVPDETYYTESQLVDP-QYENGKIIGKSPDSGHEVELVKEESYFNISKYTD 179

QY 180 RLVEFPKAHPFITPDGRNLNEMIRNFIEPGLEDIAVSRITFTWGVPYPSNPKHYVVMID 239
DB 180 RLFEFYDQPDFIQPPSRKNEMINFIKPLGLADLAVSRTSFNGVHVYPSNPKHYVVMID 239

QY 240 ALINYATALGYAQDEHGFNDFKWNGTVFMVGKDILRHFSIYWPIILLMWLLDVKLPDLRIA 299
DB 240 ALVNIYISALGYLSDDSLENKYWPADI-HLMAKEIVRFHSIIWPILLMALDLPLPKKVFA 298

QY 300 HGVFMVKDGMSKSGNTVYPEMLVERYCLDPLRYLMRNLVPGSDGTFTPEDYGRNY 359
DB 299 HGWLMDKGMSKSGKNVVDPNILDRYGLDATRYLMRELPGSDGVFTPEAFVETNF 358

QY 360 ELANDLGNLNRTVSMINKYFDGQPAYYEGVTEFDHLAEVAEKSIAADFHTHMEADVYP 419
DB 359 DLANDLGNLVNRITISMVNKYFDGELPAYQGPLHELDEEMAALETSVKSYTBSMESLQFS 418

QY 420 RALEAWTLISRNTKKIIDTAFWLLDKBALDQLASVMSHWQASIRVVVAHILEPFMET 479
DB 419 VALSTVWKFIISRTNKYIDETTFWLAKDSQKMDLGNNVAHLVENIRYAUVLLRPLFLTHA 478

QY 480 SRAVLQTQIGLE-----EVSSLENLSLADFPADVTVVAKGTPIFFPRLNMBEEIAYIKEQM 534
DB 479 PKEIPEOLINNPNQPMFEPSLSOYGV--LTESIYVTGQPKPIFPLRLDSEAEIAYIKESNQ 536

QY 535 GNKPATKEKNVPDEVLELKLNKDSEIKFEOPDKVEIRVAEVEKVSKEGSDKLLOFLDAGD 594
DB 537 --PPATEER--KEEIP---SKFOIDIKDPKRVETKAAITIDAHHVKKSKDKLKIQVDL-D 588

QY 595 GEDRQILSLGIAYYPNEQSELVGGKVIQVANLXPRMKKKYVSQGMLLSAEHDGKLTLYV 654
DB 589 SEORQIVSGIAKFY-TPDDIIIGKKVAVTNLTKPAKIMGQ-KSEGMIILSAEKDGVILTLYSL 646

RESULT 9
 S16682
 methionine-trna ligase (EC 6.1.1.10) - Bacillus stearothermophilus
 N:Alternate names: methionyl-tRNA synthetase
 C:Species: Bacillus stearothermophilus
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S16682
 R:Mechulam, Y.; Schmitt, E.; Panvert, M.; Schmitter, J.M.; Lapadat-Tapolsky, M.; Meinell
 Nucleic Acids Res. 19, 3673-3681, 1991
 A:Title: Methionyl-tRNA synthetase from Bacillus stearothermophilus: structural and func

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: S16682; MUID:91305115; PMID:1852609
A:Accession: S16682
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-649 <MEC>
A:Cross-references: UNIPROT:P23920; EMBL:X57925; NID:g39988; PIDN:CAA0999.1; PID:g39988
C:Genetics:
A:Gene: CAC2991
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 55.2%; Score 1919; DB 2; Length 649;
Best Local Similarity 55.9%; Pred. No. 45e-107;
Matches 372; Conservative 103; Mismatches 171; Indels 20; Gaps 7;
QY 1 MSEKNFYTTPIYPSGKLHIGSAVTTIACDVLAARYKELMGDYVYLTGLDEHCKIQK 60
DB 1 MEKKTFTYLTPIYPSGKLHIGSAVTTIACDVLAARYKELMGDYVYLTGLDEHCKIQK 60
QY 61 AEEAGITPOAVYDGMVGVKELWQLDLSYDKFIRTTDDYHEKVVAVQVFERLLAQDDIYL 120
DB 61 AOEKGVTFQYVDDIVAGIQELWRLKLDLSYDDFIRTTQERHKKIVEKIFARLVEQGIYL 120
QY 121 GSYGWSYSDSEFFETESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQDR 180
DB 121 GEYEGWYCTPCSEFVTERQL-----VDGNCPCDGRPVEKVEESYFFRMSKYVDR 170
QY 181 LVEFFKAPHEFTTPGRNEMLRNFIPEGLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 240
DB 171 LLQYEEENPDFIQPSRKNEMNINFIPEGLEDLAVSRFTFTWGIKVPDGRHVIYVWIDA 230
QY 241 LLNAYATAGYAOQDEHGNFDFKWNQVFMVGVKDIIRFHSIYWPILLMLMDLKLDRLLAH 300
DB 231 LANYITALGYGTDNDEKFRKYWPADV-HLVGKEIVRFHTIYWPIMLALGPLPKVFGH 289
QY 301 GWFVWKGMSKSGNVVYPEMLVRYGLDPLRYLYLNRNLPVSGDGTTPEDYVGRINYE 360
DB 290 GWLLMKDGMKSGNVVDPVMIIDRYGLDALRYLLREVFGSDGVFTPEGFERINYD 349
QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGTFEPDHLVAEAKSIADFHTHMEAVDYP 420
DB 350 LANDLGNLNRVAVLEKIFGGAIPYRGKTPFPDRDLSAREVVRVYEEAMERMEFSV 409
QY 421 ALBAVNTLISRTNKYIDETAPVWLKDQALRDQLASVMSHWQASIRVVVAHLIEBPMET 480
DB 410 ALSAVWLIIGRTNKYIDETQPVWLAKESKREELASVMAHLAESLRYTAVLLQPLTRP 469
QY 481 RAVLTQIGLEEVSSLENLSLADFP--ADVTVAKGTPIFPRLNMBEETAYIKQMEGNKP 538
DB 470 ERIFTQGISDRSLKENDSLYDFGLIPGTNVQKGEPLFPRLDICEVEVEYIKAHVQGGKP 529
QY 539 AVEKEWNPDEVELKLNKDEIKPEFDKVEIRVAEVEKVSKEVSGSKLQFLRDAGDGEDR 598
DB 530 AEAAR-----EEQAAAEBSIDDPKAVDLRVAEVVQPERMKNADKLLQLDLG-GEKR 584
QY 599 QILSGIAKYPNEQELGVKQVIVANLKPRLMKKKVVSQGMIL--SAEHGDKLTLLTVDP 658
DB 585 QVLSGIAEFYKPE-ELIGKKVTCVANLKPRLMKQ--LSQGMILAVATDDSKLYTLDPED 642
QY 659 PNGSVI 664
DB 643 PNGTKI 648

RESULT 10
A97268
methionyl-tRNA synthetase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: A97268
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97268
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <KUR>
A:Cross-references: UNIPROT:097EWS; GB:AE001437; PIDN:AAK80932.1; PID:g15026047; GSPDB:GN
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2991
C:Superfamily: methionyl-tRNA synthetase, dimer-forming

Query Match 47.2%; Score 1640.5; DB 2; Length 644;
Best Local Similarity 48.3%; Pred. No. 2e-90;
Matches 322; Conservative 125; Mismatches 193; Indels 27; Gaps 10;
QY 1 MBS-KNFTTITPIYPSGKLHIGSAVTTIACDVLAARYKELMGDYVYLTGLDEHCKIQK 59
DB 1 MSEKKFYTTPIYPSGKLHIGSAVTTIACDVLAARYKELMGDYVYLTGLDEHCKIQK 60
QY 60 KBAEAGITPOAVYDGMVGVKELWQLDLSYDKFIRTTDDYHEKVVAVQVFERLLAQDDIY 119
DB 61 IAEKXGKITPKAYVDELVAGIKDLWKNMISYDKFIRTTDEEHVKAQVQKIVKFFDNGDIY 120
QY 120 LGEYSGWYSVSDSEFFETESQLAEVFRDEAGNVTGGIAPSGHEVWVSSESYFLRLSKYQD 179
DB 121 KSAVEGWYCTPCSEFWTETQL-----VDGKPCDGRVETKBEAYFFKMSKYAD 170
QY 180 RLVEFFKAPHEFTTPGRNEMLRNFIPEGLEDLAVSRFTFTWGPVPSNPKHVVVWIDA 239
DB 171 RLKITYEDPDFIQPSRKNEMNINFIPEGLEDLAVSRFTFTWGIKVPDGRHVIYVWIDA 230
QY 240 ALNAYATAGYAOQDEHGNFDFKWNQVFMVGVKDIIRFHSIYWPILLMLMDLKLDRLLAH 299
DB 231 ALSNVITALGYGSDNDDELNKNFWPADL-HLVGKDIIRFHTIYWPIMLALDPLPKQVFG 289
QY 300 GWFVWKGMSKSGNVVYPEMLVRYGLDPLRYLYLNRNLPVSGDGTTPEDYVGRINYE 359
DB 290 HGWLLVDGGMKSGNVVDPVMIIDRYGLDALRYLLREVFGSDGVFTPEGFERINYE 349
QY 360 ELANDLGNLNRVTSMINKYFDGQIPAYVEGTFEPDHLVAEAKSIADFHTHMEAVDYP 419
DB 350 DLANDLGNLVRATAMIEKYFDGSIQPPVD-KEEIDNELIDMAISLPEKLEDDIKKLKP 408
QY 420 RALEAVNTLISRTNKYIDETAPVWLKDQALRDQLASVMSHWQASIRVVVAHLIEBPMET 479
DB 409 EALDHIWDLIKRANKYIDETTPWLAKENKARGLTVLYNLVSLRFVATLTTPFLPET 468
QY 480 SRAVLTOIGLEEVSSLENLSLADFPADVTVAAGTPIFPRLNMBEETAYIKQMEGNKPA 539
DB 469 GEKIKTQLAI-ELDTWESLSAFDGTAGTKVSGEVIFPRIDVDKIEELNKLKEQLKA 527
QY 540 VEKEMNPDEVELKLNKDEIKPEFDKVEIRVAEVEKVSKEVSGSKLQFLRDAGDGEDR 599
DB 528 TRK-----MQPLKPEISIDDDKDLRVVVKVLECEPEVKKSKLLKLVKELG-GESEQ 578
QY 600 ILSGIAKYPNEQELGVKQVIVANLKPRLMKKKVVSQGMIL--SAEHGDKLTLLTVDP 657
DB 579 VLSGISQFYKPE-DLIGKKVTVLVANLKPRLMKQ--LSQGMILAVATDDSKLYTLDPED 636
QY 658 VPNGSVI 664
DB 637 IPTGSIV 643

RESULT 11
E72297
methionine-tRNA ligase (EC 6.1.1.10) - Thermotoga maritima (strain M588)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: E72297
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence
A:Reference number: A72200; PMID:99287316; PMID:10360571
A:Accession: E72297
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-629 <ARN>
A:Cross-references: UNIPROT:O33925; GB:AE001768; GB:AE000512; NID:g4981619; PIDN:AAD3616
A:Experimental source: strain MSB8
C:Genetics:
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 39.1%; Score 1359.5; DB 2; Length 629;
Best Local Similarity 42.4%; Pred. No. 1.2e-73;
Matches 284; Conservative 127; Mismatches 204; Indels 55; Gaps 14;

QY	6	FIYTTPIYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYLTGDEHGQKIQKAEBAAG	65
DB	3	FIYTTPIYVNSPEPHIGSAYTTIADVLARYKRLMGYDVFLYLTGDEHGQKIQKAEBAAG	62
QY	66	ITPQAVVDMGVAVGKELWLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYLGEYS	125
DB	63	KDQEFCDLAKFKELWLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYLGEYS	122
QY	126	WYSVDEEFFTESQLAEVFRDEAGNVGTGGIAPSGHE-----VEWVSEESYFLRLSK	176
DB	123	WYCVPCETFWNEDE-----VIKEGERPCPECKRPVKVVEEENYFRLSK	167
QY	177	YQDLVEFFKAHPEFITPPGRNLNMLNFIPEGLEDLAVSRFTFTGVPVPSNPQHVIV	236
DB	168	YRDLKLYEHPDFVPPFRNEMLK-LLEGKDLSTRTTFKGVPMKDDPEHVIYV	226
QY	237	WIDALINYATALGYADDEHGNDFKFWGVFHMVGKDIIRFHSIYWPILLMLMDVKLPDR	296
DB	227	WVDALINYAIGYGNDE-MENKWPADL-HLIGKEINRFSIIPWMLMSVGLPLPKK	284
QY	297	LIAHGFVMDKMSKSGNVVYPEMLVERYGLDPLRYVLMNLPVSGDGTTPEDYVGR	356
DB	285	VFAHGWLTNGQKISKSLGNAIDPRFFVKRYGNDVYVYLLIRDMFGKDGDFSEERLVR	344
QY	357	INVELANDLGNLNRVTSMINKYFDGQI--PAVEGVTEFHDVLAEVAEKSIADPHTME	414
DB	345	LNSDLANDYGNLHRTAMIKYFNGRLSPSAQEG---FDSWLKERFETKDAIHEMD	401
QY	415	AVDYPALSAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQASIRVVAHLIEP	474
DB	402	SYRLTEALDKIWEFIADVNKYFNDTKPWILGK-EGNMRGLGTVLYNSLEAVEFKVALMTLP	460
QY	475	FWMTSRAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIPPRLNMEETIAYIKEQME	534
DB	461	VMPDTSEEVFRVSEFEKPSKSHLENWGLVKGSTVHGEPLFKKIDAKD-----FKVVE	516
QY	535	GNKPAVEKEWNPDEVELKLNKDEIKPEDPKVEIRVAEVEKYSKVEGSKLQFLDAGD	594
DB	517	---TVSAEQN-----AITDDRSKVDLAIAKVAEKVPNSRKULRLIIDLGT	561
QY	595	GEDRQLLSGIAYKYPNEQELVKKKVQIVANLKRPMKKYVSGMILSAEHGCKTLITV	654
DB	562	-EKRQIVAGIAEHY-RPEELVGLKIVVANLKPAKLM-GIESQGMILAAKSGDTLRLLTV	618
QY	655	DPAPVNSGVI 664	
DB	619	DGEITPGAKV 628	

RESULT 12

F81356
methionine-tRNA ligase (EC 6.1.1.10) Cj0838c [imported] - Campylobacter jejuni (strain N
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: F81356
R:Packhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: AB1250; PMID:20150912; PMID:10688204
A:Accession: F81356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-628 <PAR>
A:Cross-references: UNIPROT:Q9P85; GB:AL1139076; GB:AL111168; NID:g6968128; PIDN:CAB7310;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
C:Superfamily: methionyl-tRNA synthetase, dimer-forming
C:Keywords: ligase

Query Match 34.2%; Score 1189.5; DB 2; Length 628;
Best Local Similarity 33.2%; Pred. No. 1.8e-63;
Matches 266; Conservative 132; Mismatches 209; Indels 71; Gaps 18;

QY	7	YITTPYPSGKLHIGSAYTTIACDVLARYKRLMGYDVFLYLTGDEHGQKIQKAEBAAG	66
DB	3	YITTPYVNDVPHLGAYTTIADTLARFVRLQGHEHTRFLTGDEHGQKIEEAAKLNS	62
QY	67	TPQAVVDMGVAVGKELWLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYLGEYS	126
DB	63	TPQAVVDMGVAVGKELWLLDISYDKFIRTTDDVHEKVAQVFERLLAQDDIYLGEYS	122
QY	127	YSVDEEFFTESQLAEVFRDEAGNVGTGGIAPS-GHEVEWVSEESYFLRLSKYQDLVEFF	185
DB	123	YICSESEFFTSQJ-----INDCSPDCKQKTRILKESYFFKLKQKILQWY	172
QY	186	KAHPEFITPPGRNLNMLNFIPEGLEDLAVSRFTFTGVPVPP-----SNPKHVYVYVWDALL	242
DB	173	E-KDPIPKNKNKELL-NFQVGLKDLSTRTTFKGVPMKDDPEHVIYV	230
QY	243	NYATGVAQDEHGNDFKFWGVFHMVGKDIIRFHSIYWPILLMLMDVKLPDLIAHGW	302
DB	231	IYVSSLDL--QNGENAKFPFAHV-HLVGKDIIRFHAHYWPAFLMSVDLPLPKFIGA	287
QY	303	FVMKDGMSKSGNVVYPEMLVERYGLDPLRYVLMNLPVSGDGTTPEDYVGRINYLEA	362
DB	288	WTKGERMSKSGNVVYPEMLVERYGLDPLRYVLMNLPVSGDGTTPEDYVGRINYLEA	347
QY	363	NDLGNLNRVTSMINKYFDGQI-PAYVEGVTEFHDVLAEVAEKSIADPHTMEAVDYPRAL	422
DB	348	NEFGNLLNRITIGMSTKYSQGNISK--EGVLKFYNAELNQAHEHLNLAVEFLENLQCNRYL	405
QY	423	EAVWTLISRTNKYIDETAPWVLDKDEALRDQLASVMSHQASIRV-----VAHLIE	473
DB	406	EELFKALSVANLAIKSKYEPHSLIKEN-----KHEQANALVALCANILAKTSLLS	455
QY	474	PFMTSRAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIPPRLNMEETIAYIKEQME	527
DB	456	PTLPKSKQKVALANFETISSANVTKMILDN-ELLDFKANPC-----EALPKV-----EKA	505
QY	528	YIKQMEGNKPAVEKEWNPDEVELKLNKDEIKPEDPKVEIRVAEVEKYSKVEGSKLQ	587
DB	506	LLKQEI---KEEPPKEESP-----KIKIDDFAKIEIKVAKVLDCCNIGSEKLLK	552
QY	588	FRLDAGDEDRQLLSGIAYKYPNEQELVKKKVQIVANLKRPMKKYVSGMILSAEHG	647
DB	553	FQELDDKEIRVQLSGIAKYY-KASDLIGKQVCVISNLKKAIF-GHESDGMILSAKSGD	610
QY	648	KLTLTVDPAPVNSGVIG 665	
DB	611	KLVLIAPEQLVQNGSLVG 628	

RESULT 13

C82922
methionyl-tRNA synthetase UUI97 [imported] - Ureaplasma urealyticum

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: A64572

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-650 <TOM>

A:Cross-references: UNIPROT:P56127; GB:AB000557; GB:AB000511; NID:g2313514; PIDN:AAD0748

C:Superfamily: methionyl-tRNA synthetase, dimer-forming

C:Keywords: ligase

```
Query Match      33.0%; Score 1148.5; DB 2; Length 650;
Best Local Similarity 39.6%; Pred. No. 5.3e-61;
Matches 268; Conservative 120; Mismatches 236; Indels 53; Gaps 16;

Qy      8 ITTPIYVPSGKLIHGSAVTTIACDVLARYKELMGVDVYLTGLDGHGKIQQKAEERAGIT 67
Db      6 ITTPIYVNDIPHCHAYTHLIADTLKYYTLQGEVFFLTGDEHGCKIEQEARLNQS 65
Qy      68 PQAYVDMAGVVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGXYSGWY 127
Db      66 PKAYADGISITFKQWDFNFNDYDGFIRTTDSEHQKCVQNAFIMPEKGDYKGYSGY 125
Qy      128 SVSDEEFFTSQLAEVRDEAGNVGCIAPSG-HEVEWVSEESYFRLSKYQDRLYVEFFK 186
Db      126 CVSCESYCAISK-ADNTNDKV-----LCPDCURETTLLSEESYFFRLSAYEKPLDFYA 178
Qy      187 AHPEFITPDGRLNEMLRNFIEPGLIEDLAVSRITFTMGVPVP---SNPKHVYVYVWIDALLN 243
Db      179 KNPEAILPVYRKNE-VTSFIEQGLLDLSITRTSPFWGIPLPKKNWDPKHVVYVWLDALLN 237
Qy      244 YATALGYAODEHGNFDFWNGTVFHMVGKDIILPHSYIWFILLMLDLVKLPDRLLAHGWF 303
Db      238 YASALGYNLNDLNQMAHF--ECARHIVGKDIRFRHAIYPAFLMSLNLPLFKQLCVHGW 295
Qy      304 VMKDGKMSKSGNVVYVPEMLVERGYLDPLRYLMLNLPVGSDDGTFTPTEDYVGEINVELAN 363
Db      296 TIEGVNMSKSLGNVLDQAQKAMEYGIIEELRYFILLREVFGQDGDGFSKKALIERINANLNN 355
Qy      364 DLGNLNRVTSMINKYFD----QIPAYVEGVTEFDHVLAEVAEKSIADFTHTMEAVDY 418
Db      356 DLGNLNRLLIGMAKYYFNHSLKSTKITAYYSKELEKVHQILDNANSFVPMQQLH----- 409
Qy      419 PRALAEVTLISRTNKYIDETAPWLDKDEALRDQDLASVMHWOASIRVVAHLIEPFM-- 476
Db      410 -KALEELFNYYDFLNKLIKAKEEPWLVKKNNE-SEKLEALLSLANALQSSFLLYAFMPK 467
Qy      477 --METSRAVLTOIGLE-----EVSSLENLSLADFPADVTVAKGTPIPPLANWEEIAY 528
Db      468 SAVKLANAFNTEITPDNYERFFKAKKLQMILOD-----TEPLFSKM---EKIEK 514
Qy      529 IKQMEGNKFAVEKNPDEVELKLNKDE-IKPFDFKVEIRVAEYKSVKVEGSKLLQ 587
Db      515 TERAGEASPEKNEKEKDAKAPLQENYIGIEDPKKVEIKVGLIKEAQRIEKNKLLR 574
Qy      588 FRLDAGGEDRQILSGIAYKYPNEQELVGKVOIVANLKPOMKXYVSGMILSAHDG 647
Db      575 LKVDLGEGRLRQIISGIALDYEPF-SLVGMQVCVAVNLKPAKLME-MSEGMI LAVRDS 632
Qy      648 KLTLLTVDPAPVNGSVI 664
Db      633 NLALISPTREKIAGSLI 649
```

Search completed: December 17, 2004, 22:00:18

Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2004, 21:42:30 ; Search time 200 Seconds
(without alignments)
193.121 Million cell updates/sec

Title: US-10-606-607-2
Perfect score: 3479
Sequence: 1 MSEKNFYITPIYPSGKLH.....DGKLLTLLTVDPAVNGSVIG 665

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3446	99.1	665	1 SYM_STRPN	Q9716 streptococ
2	2883	82.9	669	2 Q8DSW8	Q8DSW8 streptococ
3	2880.5	82.8	665	1 SYM_STRP3	P59080 streptococ
4	2879.5	82.8	665	1 SYM_STRP5	Q9A178 streptococ
5	2868.5	82.5	665	1 SYM_STRP8	Q8P298 streptococ
6	2732.5	78.5	665	2 Q8DYD0	Q8DYD0 streptococ
7	2732.5	78.5	665	2 Q8E3Z3	Q8E3Z3 streptococ
8	2392.5	68.8	682	1 SYM_LACLA	Q9CHE0 lactococcu
9	2288	65.8	669	1 SYM_ENTFA	Q837B3 enterococ
10	2230.5	64.1	673	1 SYM_OCEIH	P59079 oceanobacil
11	2201	63.3	683	1 SYM_LACPL	Q88297 lactobacil
12	2114	60.8	665	2 Q72FN6	Q724N6 listeria mo
13	2114	60.8	665	2 AAT02975	AAT02975 listeria
14	2113	60.7	664	1 SYM_LISMO	Q8YAF2 listeria mo
15	2107	60.6	664	1 SYM_LISIN	Q92190 listeria in
16	1990	57.2	658	2 Q74LI3	Q74LI3 lactobacil
17	1990	57.2	658	2 AAS08184	AAS08184 lactobacil
18	1963.5	56.4	664	1 SYM_BACSU	P37465 bacillus su
19	1953.5	56.2	660	1 SYM_BACHD	Q9KGK8 bacillus ha
20	1951.5	56.1	660	2 Q73FH0	Q73FH0 bacillus ce
21	1951.5	56.1	660	2 AAS38972	AAS38972 bacillus
22	1939.5	55.7	660	2 Q6HFX8	Q6HFX8 bacillus th
23	1936.5	55.7	660	1 SYM_BACOR	Q81JA8 bacillus ce
24	1936	55.6	657	2 Q6GJI1	Q6GJI1 staphylococ
25	1935.5	55.6	660	1 SYM_BACAN	Q81W03 bacillus an
26	1935.5	55.6	660	2 AAT79115	AAT79115 bacillus
27	1931	55.5	657	1 SYM_STRAM	Q99WB3 staphylococ
28	1929	55.4	657	1 SYM_STRAM	Q8NY00 staphylococ
29	1929	55.4	657	2 Q6GSZ8	Q6GSZ8 staphylococ
30	1922.5	55.3	656	1 SYM_STRAP	Q8CQU3 staphylococ
31	1919	55.2	643	1 SYM_BACST	P23920 bacillus st

32	1701	48.9	638	1 SYM_THETN	Q8RDD1 thermoanaer
33	1664	47.8	641	1 SYM_CLOTE	Q899D9 clostridium
34	1651.5	47.5	645	1 SYM_CLOPE	Q8XNG1 clostridium
35	1640.5	47.2	644	1 SYM_CLOAB	Q97EWS clostridium
36	1499	43.1	636	2 Q7P484	Q7P484 fusobacteri
37	1484	42.7	636	1 SYM_FUSNN	Q8RE57 fusobacteri
38	1421.5	40.9	418	2 Q7WSP6	Q7WSP6 bacillus ce
39	1417.5	40.7	416	2 Q7WSQ0	Q7WSQ0 bacillus ce
40	1407.5	40.5	416	2 Q7WSR5	Q7WSR5 bacillus ce
41	1387.5	39.9	411	2 Q7WSR7	Q7WSR7 bacillus ce
42	1382.5	39.7	410	2 Q7WSP8	Q7WSP8 bacillus ce
43	1359.5	39.1	629	1 SYM_THEMA	Q33925 thermotoga
44	1340.5	38.5	400	2 Q7WSQ1	Q7WSQ1 bacillus ce
45	1329.5	38.2	397	2 Q7WSP9	Q7WSP9 bacillus ce

ALIGNMENTS

RESULT 1

ID	SYM_STRPN	STANDARD;	PRT;	665 AA.
AC	Q97RL6			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).			
DE	Name=metG; Synonyms=metS; OrderedLocusNames=SP0788, spr0696;			
OS	Streptococcus pneumoniae, and			
OS	Streptococcus pneumoniae (strain ATCC BAA-255 / R6).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313, 171101;			
RN	[1]_TaxID=1313, 171101;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC BAA-334 / TIGR4;			
RX	MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;			
RA	Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,			
RA	Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,			
RA	Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,			
RA	Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,			
RA	Holtzapple E.K., Khouri H.M., Wolf A.M., Uterback T.R., Hansen C.L.,			
RA	McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,			
RA	Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,			
RA	Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;			
RT	"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae."			
RL	Science 293:498-506(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC BAA-255 / R6;			
RX	MEDLINE=21429245; PubMed=11544234;			
RX	DOI=10.1128/JB.193.19.5709-5717.2001;			
RA	Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,			
RA	DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geisinger C.,			
RA	Gilmour R., Glass J.S., Kuoja H., Kraft A.R., Lagace R.E.,			
RA	LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushina P.,			
RA	McAhen S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,			
RA	Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,			
RA	Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,			
RA	Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,			
RA	Glass J.I.;			
RT	"Genome of the bacterium Streptococcus pneumoniae strain R6.;"			
RL	J. Bacteriol. 183:5709-5717(2001).			
CC	-!- FUNCTION: Is required not only for elongation of protein synthesis			
CC	but also for the initiation of all mRNA translation through			
CC	initiator tRNA(fMet) aminoacylation (By similarity).			
CC	-!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +			
CC	diphosphate + L-methionyl-tRNA(Met).			
CC	-!- SUBUNIT: Homodimer (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.			

CC MetG subfamily 2B.
 CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE007385; RAK74925.1; -;
 CC EMBL; AE008446; AAK99500.1; ALT_INIT.
 CC FIR; D95091; D95091.
 CC HSSP; Q9V011; 1MKH.
 CC TIGR; SP0788; -;
 CC HAMAP; MF 01228; fused; 1.
 CC InterPro; IPR004495; MetG Cterm.
 CC InterPro; IPR008224; MetRS dimerising.
 CC InterPro; IPR002300; tRNA-synt_1a.
 CC InterPro; IPR001412; tRNA-synt_1.
 CC InterPro; IPR002304; tRNA-synt_met.
 CC InterPro; IPR009080; tRNAasn_1a_bind.
 CC InterPro; IPR002547; tRNA_bind.
 CC Pfam; PF01333; tRNA-synt_1; 1.
 CC Pfam; PF01588; tRNA_bind_1.
 CC PIRSP; PIRSP001528; MetRS dimerising; 1.
 CC PRINTS; PR01041; TRNASYNTHMET.
 CC TIGRFAMs; TIGR00398; metG; 1.
 CC TIGRFAMs; TIGR00178; AA tRNA_LIGASE_1; FALSE_NEG.
 CC PROSITE; PS00896; TRBD; 1.
 CC AMINOACYL-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Protein biosynthesis; RNA-binding; tRNA-binding.
 KW SITE 13 23 "HIGH" region.
 FT SITE 309 313 "KMSKS" region.
 FT DOMAIN 562 665 tRNA-binding.
 FT BINDING 312 312 ATP (By similarity).
 SQ SEQUENCE 665 AA; 75651 MW; 7FAB60A5CC49A56D CRC64;
 Query Match 99.1%; Score 3446; DB 1; Length 665;
 Best Local Similarity 99.2%; Pred. No. 3.5e-198;
 Matches 660; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSEKFIYITPIYPSGKLHIGSAVTTTACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 60
 DB 1 MSEKFIYITPIYPSGKLHIGSAVTTTACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 60
 QY 61 ABEAGITPAYDGMAGVKELWQLDTSYDKFIITDDYHEKVAQVFERLLAQDDIYL 120
 DB 61 ABEAGITPAYDGMAGVKELWQLDTSYDKFIITDDYHEKVAQVFERLLAQDDIYL 120
 QY 121 GEYSGWYSDDEFFTESQAEVFRDEAGNVGGIAPSGHEVVEWVSSESYFLRLSKYQDR 180
 DB 121 GEYSGWYSDDEFFTESQAEVFRDEAGNVGGIAPSGHEVVEWVSSESYFLRLSKYQDR 180
 QY 181 LVEFFKAHPEFITPGRNLEMLNFIPEGLEDLAVSRITFTTGVVPVSPKRVVYVWIDA 240
 DB 181 LVEFFKAHPEFITPGRNLEMLNFIPEGLEDLAVSRITFTTGVVPVSPKRVVYVWIDA 240
 QY 241 LNNYATGACADEGNFDFKNGVTGVHMGDKILRFHSIYWPILLMLDKLPDLRLIAH 300
 DB 241 LNNYATGACADEGNFDFKNGVTGVHMGDKILRFHSIYWPILLMLDKLPDLRLIAH 300
 QY 301 GFWFMKDGKMSKGNVVPYEMLVRYGLDPLRYLYLMRNLVPGSDGTFTPEDYVGRINYE 360
 DB 301 GFWFMKDGKMSKGNVVPYEMLVRYGLDPLRYLYLMRNLVPGSDGTFTPEDYVGRINYE 360
 QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEFDHVAEVAEKSIADPHTHMEADVPR 420
 DB 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEFDHVAEVAEKSIADPHTHMEADVPR 420
 QY 421 ALEAVVTLISRTNKYIDETAPVWLAKDEALRDLQSLASVMSHWSQASIRVVAHLIEPFPMETS 480

DB 421 ALEAVVTLISRTNKYIDETAPVWLAKDEALRDLQSLASVMSHWSQASIRVVAHLIEPFPMETS 480
 QY 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTFIPFLANMEEIAY:KEQMEGNKPAV 540
 DB 481 RAVLTQIGLEEVSSLENLSLADFPADVTVVAKGTFIPFLANMEEIAY:KEQMEGNKPAV 540
 QY 541 EKWNVDEVELKLNKDEIKFEDFDKVEIRVAEVEKVEGSDKLLQPLDAGDGEDROI 600
 DB 541 EKWNVDEVELKLNKDEIKFEDFDKVEIRVAEVEKVEGSDKLLQPLDAGDGEDROI 600
 QY 601 LSGIAKYYPNEQELVGKQVIVANLKPRKMKKYYSQGMILSAEHDGKLTLLTVPDAVEN 660
 DB 601 LSGIAKYYPNEQELVGKQVIVANLKPRKMKKYYSQGMILSAEHDGKLTLLTVPDAVEN 660
 QY 661 GSVIG 665
 DB 661 GSVIG 665
 RESULT 2
 QSDSW8 ID QSDSW8 PRELIMINARY; PRT; 669 AA.
 AC QSDSW8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DB Putative methionyl-tRNA synthetase (EC 6.1.1.10).
 GN Names: OrderedLocustNames=SMU.1639;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE014994; AAN59279.1; -;
 DR HSSP; Q9V011; 1MKH.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004825; F:methionine-tRNA ligase activity; IEA.
 DR GO; GO:0000049; F:tRNA binding; IEA.
 DR GO; GO:0006431; P:methionyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNAasn_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF01333; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind_1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRFAMs; TIGR00398; metG; 1.
 DR TIGRFAMs; TIGR00178; AA tRNA_LIGASE_1; FALSE_NEG.
 DR PROSITE; PS00896; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
 SQ SEQUENCE 669 AA; 75635 MW; 10B0B3ACES1DDCEE CRC64;
 Query Match 82.9%; Score 2883; DB 2; Length 669;
 Best Local Similarity 82.4%; Pred. No. 2.1e-164;
 Matches 553; Conservative 46; Mismatches 64; Indels 8; Gaps 5;
 QY 1 MSEKNFYITPIYPSGKLHIGSAVTTTACDVLARYKRLMGVDVFLYGLDGHGQKIQQ 59

Db 1 MTEKQPFYITPIYPSGKLHIGSAVTTIACDVLARYKEMMNDVFYLTGLDEHQKIQ 60
QY 60 KAEAGITPQAVYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 119
Db 61 KSEAGITPQAVYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIY 120
QY 120 LGEYSWYSVDSDEFFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYOD 179
Db 121 LGEYSWYSVDSDEFFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYAD 180
QY 180 RLVEPFKAPERTTPDGRNEMLRNIEPGLDVLAVSRITFTWGVPPVPSNPKHVYVWID 239
Db 181 RLVDFFHAPDFIOPGRNEMLRNIEPGLDVLAVSRITFTWGVPPVPSNPKHVYVWID 240
QY 240 ALLNYATAGYADDEHGNFKPWNVTFFHMGVDKILRPHSIYWPILLMLDKVLPDLRLFA 299
Db 241 ALLNYATAGYADDEHGNFKPWNVTFFHMGVDKILRPHSIYWPILLMLDKVLPDLRLFA 300
QY 300 HGFVWKDGMKSGKNVYYPENLVRYGLDPLRYLXNLPVSGSDGTTTPEDYVGRINY 359
Db 301 HGFVWKDGMKSGKNVYYPENLVRYGLDPLRYLXNLPVSGSDGTTTPEDYVGRINY 360
QY 360 ELANDLGNLNRVSMINKYFDQIPAYVEGVTFEPDHLVAEAKSIADFTHTMEADVYP 419
Db 361 ELANDLGNLNRVSMINKYFDQIPAYVEGVTFEPDHLVAEAKSIADFTHTMEADVYP 420
QY 420 RALEAVWTLISRTNKYIDETAPWLDKDEALRDLASVMSHQASTRVVAHIEPFPMBET 479
Db 421 RALEAVWTLISRTNKYIDETAPWLDKDEALRDLASVMSHQASTRVVAHIEPFPMBET 480
QY 480 SRAVLQIGLEEVSSLENISLADFPADVTVAAKGTPIFPLNMEEBIAIYKCEMEGNKPA 539
Db 481 SRAVLQIGLEEVSSLENISLADFPADVTVAAKGTPIFPLNMEEBIAIYKCEMEGNKPA 539
QY 540 V-----EKNWPNDRVELKANKDEIKFEDFKVEIRVAEVEKVSQSGDKLLOFRLDAGDG 595
Db 540 V-----EKNWPNDRVELKANKDEIKFEDFKVEIRVAEVEKVSQSGDKLLOFRLDAGDG 599
QY 596 EDQQLISGIAKYPNQELVGVKVOIVANLKPKNMKVYVSGQMILSAEH-DGKLTLLTV 654
Db 600 EDQQLISGIAKYPNQELVGVKVOIVANLKPKNMKVYVSGQMILSAEH-DGKLTLLTV 658
QY 655 DPVAPNGSVITG 665
Db 659 DSSVPNGSQIG 669

RESULT 3
SYM_STRP3 STANDARD; PRT; 665 AA.
AC P59080;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (MetRS).
GN Name=metG; Synonyms=metS; OrderedLocusNames=SpyM3_0300, Sps1557;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RA MEDLINE=22133809; PubMed=12122206; DOI=10.1073/pnas.152298499;
RX Bares S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
phage-encoded toxins, the high-virulence phenotype, and clone
emergence."
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).

(2)
SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
large-scale genomic rearrangement in invasive strains and new insights
into phage evolution."
RL Genome Res. 13:1042-1055 (2003).
CC -!- FUNCTION: Is required not only for elongation of protein synthesis
but also for the initiation of all mRNA translation through
initiator tRNA (Met) aminoacylation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA (Met) = AMP +
diphosphate + L-methionyl-tRNA (Met).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2b.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; AE014142; AAW78907.1; -
DR EMBL; AP005145; BAC64652.1; ALT_INIT.
DR HSSP; P23395; 1A8H.
DR HAVAP; MF_01428; fused; 1.
DR InterPro; IPR004495; MetG_Cterm.
DR InterPro; IPR008224; MetRS_dimerising.
DR InterPro; IPR002300; tRNA-synt 1a.
DR InterPro; IPR001412; tRNA-synt 1.
DR InterPro; IPR002304; tRNA-synt_met.
DR InterPro; IPR009080; tRNA-synt_la_bind.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF01333; tRNA-synt 1; 1.
DR Pfam; PF01588; tRNA_bind; 1.
DR PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR PRINTS; PR01041; TRNASYNTHMET.
DR TIGRFAMs; TIGR00398; metG; 1.
DR TIGRFAMs; TIGR00399; metG_Cterm; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR PROSITE; PS50886; TRBD; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis; RNA-binding; tRNA-binding.
FT SITE 12 22 "HIGH" region.
FT SITE 308 312 "KMSKS" region.
FT DOMAIN 562 665 tRNA-binding.
FT BINDING 311 311 ATP (By similarity).
SQ SEQUENCE 665 AA; 74884 MW; AD5140C0C91BCFE6 CRC64;
Query Match 82.8%; Score 2880.5; DB 1; Length 665;
Best Local Similarity 82.3%; Pred. No. 2.9e-164;
Matches 547; Conservative 51; Mismatches 64; Indels 3; Gaps 2;
QY 3 EKNFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVYLTGLDEHQKIQKAE 62
Db 2 KKFYITPIYPSGKLHIGSAVTTIACDVLARYKRLMGHEVYLTGLDEHQKIQTKAK 61
QY 63 RAGITPQAVYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGE 122
Db 62 RAGITPQAVYDGMAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQDDIYLGE 121
QY 123 YSGWYSVDSDEFFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYQDRLV 182
Db 122 YSGWYSVDSDEFFTESQAEVFRDEAGNVTGGIAPSGHEVWVSEESYFLRLSKYADRLV 181

QY 541 EKEWNPDEVELKLNKDBIKFEDEPKVEIRVAEVEKSKVEGSKLLQFRILDAGDGEDROI 600
Db 541 EKEWNPDEVALKEKDVITETFDVAEIRVAEVEKSKVEGSKLLRFRVDAGDGDROI 600
QY 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVEN 660
Db 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVEN 660
QY 661 GSVIG 665
Db 661 GSIIG 665

RESULT 5
SYM_STRPB STANDARD; PRT; 665 AA.
AC QSP298;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 03-JUL-2004 (Rel. 44, Last annotation update)
DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE (MetRS)
GN Name=metG; Synonyms=metS; OrderedLocusNames=spyM18_0468;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=MGAS232 / Serotype M18;
RA PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.D., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
CC -/- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA(fMet) aminoacylation (By similarity).
CC -/- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -/- SUBUNIT: Homodimer (By similarity).
CC -/- SUBCELLULAR LOCATION: Cytoplasmic.
CC -/- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -/- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009986; AL97197.1; -
DR HSP; P23395; IAS8.
DR HAMAP; MF 01228; fused; 1.
DR InterPro; IPR004495; MetG Cterm.
DR InterPro; IPR008224; MetRS dimerising.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002304; tRNA-synt_met.
DR InterPro; IPR003080; tRNA-synt_1a_bind.
DR Pfam; PF00133; tRNA-synt_1.
DR Pfam; PF01588; tRNA_bind; 1.
DR PIRSF; PIRSF001528; MetRS dimerising; 1.
DR PRINTS; PR01041; TRNASYNTHET.
DR TIGRFAMs; TIGR00398; metG; 1.

TIGRFAMs; TIGR00399; metG_C term; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR PROSITE; PS00886; TRSD; 1.
KW Aminoacyl-tRNA synthetases; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis; RNA-binding; tRNA-binding.
FT SITE 12 22 "HIGH" region.
FT SITE 308 312 "KMSKS" region.
FT DOMAIN 562 665 tRNA-binding.
FT BINDING 311 311 ATP (By similarity).
SQ SEQUENCE 665 AA; 74869 MW; 4451CF7CF2A2D8 CRC64;
Query Match 82.5%; Score 2868.5; DB 1; Length 665;
Best Local Similarity 81.7%; Pred. No. 1.5e-163;
Matches 543; Conservative 55; Mismatches 64; Indels 3; Gaps 2;
QY 3 EKNEYITTPYVPSGKLHIGAYTTIACDVLYARKLMGYDVFYLTGLDEHGKIQKAE 62
Db 2 KAPFYITTPYVPSGKLHIGAYTTIACDVLYARKLMGYDVFYLTGLDEHGKIQKAE 61
QY 63 EAGITPQAYVDGMVGVKELMQLDLSYDKFIRITDDYHEKVAAQVFFRLIAQDDIYLGE 122
Db 62 EAGISPTQYVDNMAKDVKALMQLDLSYDTFIRITDDYHEEVAAPFEKLLAQDDIYLGE 121
QY 123 YSGWYSVSDERFFTESQLAEVFRDEAGNVGTGIIAPSGHEVEWSEESYFLRLSKYQDLRLV 182
Db 122 YSGWYSVSDERFFTESQLAEVFRDEAGNVGTGIIAPSGHEVEWSEESYFLRLSKYQDLRLV 181
QY 183 EFKAHPEFITPDGRLNEMLRNFIEPGLDLAVSGRTFTWGPVPSNPKHVVYVWIDALL 242
Db 182 AFFKERPDFIQPDGEMNEWKNFIEPGLDLAVSGRTFTWGPVPSNPKHVVYVWIDALL 241
QY 243 NVATAGVAQDEHGNFDFKFWNGTVFHMVKDILRPHSYWFIILLMLDLKPLDRLIAHGW 302
Db 242 NVATAGYGOANHANFDFKFWNGTVFHMVKDILRPHSYWFIILLMLDLKPLDRLIAHGW 301
QY 303 FVKDKGMSKSGNVVYVPEMLVERVGLDPLRYLMRNLPVSGDGTFTFEDVVGRIYELA 362
Db 302 FVKDKGMSKSGNVVYVPEMLVERVGLDPLRYLMRNLPVSGDGTFTFEDVVGRIYELA 361
QY 363 NDLGNLLNRTVSMINKYFDGQIPAYVEGVTEFDHVLAEVAKESIADPHTHMEAVDYPRL 422
Db 362 NDLGNLLNRTVAMINKYFDGTPAYVNDNGTAFDADLSQLDAQLADYHKHMEAVDYPRL 421
QY 423 EAVWTLISRTNKYIDETAPWLDKDEALRDOLASVMHMQASIRVVAHLISPMMETSA 482
Db 422 EAVWTLIARTNKYIDETAPWVLAKEDGDKQALASVMAHSLRVAHVHVIQPFMMETSA 481
QY 483 VLTQIGLEEVSSLENLSLADFPADVTVVAKGTPIPPRLNMEEEIAYIKQMEGNKPAY-- 540
Db 482 IMAQLGLAPVSDLSLALADFPANTKVAKGTPIPPRLDMEAEIDYIKAQMGDSSAISQ 540
QY 541 EKEWNPDEVELKLNKDBIKFEDEPKVEIRVAEVEKSKVEGSKLLQFRILDAGDGEDROI 600
Db 541 EKEWNPDEVALKEKDVITETFDVAEIRVAEVEKSKVEGSKLLRFRVDAGDGDROI 600
QY 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVEN 660
Db 601 LSGIAKYYPNEQELVGKKQIVANLKPRQMKKYVSGMILSAEHGDKLTLLTVDPAVEN 660
QY 661 GSVIG 665
Db 661 GSIIG 665

RESULT 6
QSDYDO PRELIMINARY; PRT; 665 AA.
AC QSDYDO;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Methionyl-tRNA synthetase.
GN Name=metG; OrderedLocusNames=SAG1557;

OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=2222298; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Tettelin H., Wasiņani V., Cielesiewicz M.J., Eisen J.A., Peterson S.N.,
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
 RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
 DR EMBL; AEO14263; AAN00423.1; -.
 DR HSSP; P23395; 1A8H.
 DR TIGR; SAG1557; -.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0005224; F:ATP binding; IEA.
 DR GO; GO:0004825; F:methionine-tRNA ligase activity; IEA.
 DR GO; GO:0000049; F:RNA binding; IEA.
 DR GO; GO:0006431; P:methionyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR002300; tRNA-synt_la.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-synt_la_bind.
 DR InterPro; IPR002547; tRNA bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA-synt_1; 1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRfams; TIGR00398; metG; 1.
 DR TIGRfams; TIGR00399; metG Cterm.
 DR PROSITE; PS00886; TRBD; 1.
 DR Aminoacyl-tRNA synthetase; Complete proteome.
 KW SEQUENCE 665 AA; 75589 MW; 7C54FE064D053C1D CRC64;
 Query Match 78.5%; Score 2732.5; DB 2; Length 665;
 Best Local Similarity 78.2%; Pred. No. 2.2e-155;
 Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;
 QY 3 EKRFYITRPIYPSGKLHIGSAVTTIACDVLARYKELMGYDYVYLTGLDGHGQKIQQKAE 62
 DB 4 KKSFYITRPIYPSGKLHIGSAVTTIACDVLARYKELMGYDYVYLTGLDGHGQKIQQKAE 63
 QY 63 EAGITQAYVDGMAVGKELWQLLDISYDKFRTTDDYHEKVAQVFERLLAODDIYLGE 122
 DB 64 EAGITQAYVDGMAVGKELWQLLDISYDKFRTTDDYHEKVAQVFERLLAODDIYLGE 123
 QY 123 YSGWYSVDEEFTESQLAEPFRDAGNTGGIAPSGHEVWVSESYFLRLSKYQDRLV 182
 DB 124 YTGWYSVDEEFTESQLAEPFRDAGNTGGIAPSGHEVWVSESYFLRLSKYQDRLV 183
 QY 183 EFPKAPHEFITDGRNEMLNRFIEPGLDLEAVSRFTTFTWGVPSNPKHVVYVWIDALL 242
 DB 184 AYAHEPFIQDGRNEMLNRFIEPGLDLEAVSRFTTFTWGVPSNPKHVVYVWIDALL 243
 QY 243 NYATAGLYA-QDEHGDFKFWNGTVFHWVKDILRHSIYWPILLMLDVKLPDLIAHG 301
 DB 244 NYTALGYGWSDDLQYHKFWPADI-HMGKILRHSIYWPILLMLDVKLPDLIAHG 302
 QY 302 WFWKDGKMSKGNVYVPEMLVERVGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 361
 DB 303 WFWKDGKMSKGNVYVPEMLVERVGLDPLRYLMLNLPVSGDGTFTPDYVGRINVEL 362
 QY 362 ANDLGNLARTVSMINKYFDGQIPAVYVGVTEPDHVLAEVASKIADPHTHVEAYDYPRA 421

Db 363 ANDLGNLARTVSMINKYFDGQIPAVYVGVTEPDHVLAEVASKIADPHTHVEAYDYPRA 421
 QY 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVWSHQASIRVVAHLIEPFFMMETSR 481
 Db 422 LEAVWTLISRTNKYIDETAPWLDKDEALRDOLASVWSHQASIRVVAHLIEPFFMMETSD 481
 QY 482 AVTQTGLEEVSSLENLSLADFPADVTVVAKGTPFPRLNMEETAYIKQEGKHPAYE 541
 Db 482 AIMEQLGLAGTDFLEKLTADLPFEGVRVAVKGSPIPRLDMEDEITYIKQEGKHPAYE 541
 QY 542 KENPDEVVELKLNKDEKPEDPKVEIRVAEYKVEKSGSKLLQFLRDAGDGEDRQIL 601
 Db 542 KENPDEVVELKLNKDEKPEDPKVEIRVAEYKVEKSGSKLLQFLRDAGDGEDRQIL 601
 QY 602 SGIATYPNEQELVGVKQIVANLKPRKMMKYYVSGMILSAEHDGKLTLLTVDPAPVNG 661
 Db 602 SGIATYPNEQELVGVKQIVANLKPRKMMKYYVSGMILSAEHDGKLTLLTVDPAPVNG 661
 QY 662 SVTG 665
 Db 662 SIIG 665
 RESULT 7
 Q8E3Z3 PRELIMINARY; PRT; 665 AA.
 ID Q8E3Z3
 AC Q8E3Z3
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Methionyl-tRNA synthetase.
 GN Name=metS; OrderedLocustNames=gbs1611;
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease.";
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766852; CAD47270.1; -.
 DR HSSP; P23395; 1A8H.
 DR SagalList; gbs1611; -.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0005224; F:ATP binding; IEA.
 DR GO; GO:0004825; F:methionine-tRNA ligase activity; IEA.
 DR GO; GO:0000049; P:RNA binding; IEA.
 DR GO; GO:0006431; P:methionyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR002300; tRNA-synt_la.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR003080; tRNA-synt_la_bind.
 DR InterPro; IPR002547; tRNA bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA-synt_1; 1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRfams; TIGR00398; metG; 1.
 DR TIGRfams; TIGR00399; metG Cterm; 1.
 DR PROSITE; PS00886; TRBD; 1.
 DR Aminoacyl-tRNA synthetase; Complete proteome.
 KW SEQUENCE 665 AA; 75589 MW; 7C54FE064D053C1D CRC64;
 Query Match 78.5%; Score 2732.5; DB 2; Length 665;
 Best Local Similarity 78.2%; Pred. No. 2.2e-155;
 Matches 519; Conservative 61; Mismatches 81; Indels 3; Gaps 3;

QY 3 EKVFYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGCKIQKAE 62
DB 4 KKSFYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGCKIQKAE 63
QY 63 EAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDDIYLGE 122
DB 64 EAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDDIYLGE 123
QY 123 YSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRIV 182
DB 124 YTCWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRIV 183
QY 183 EFKAHPEPTTPGRNEMLRNIEPGLDLAVSRFTTGWGVPVSPNKHVVVWIDALL 242
DB 184 AYAEHPEFTTPGRNEMLRNIEPGLDLAVSRFTTGWGVPVSPNKHVVVWIDALL 243
QY 243 NYATALGYA-QDEHGNFDFKFWGTVFHWGKQILRPHSYWPIILLMMLDVKLPDRILAH 301
DB 244 NYISALGYGWSDDLQYHKEWPADI-HMIGKQILRPHSYWPIILLMMLDVKLPDRILAH 302
QY 302 WFNKQGMKSKSGNVVYEMLVRYGLDPLRYILMRNLVPGSDGFTTTPEDYVGRINYL 361
DB 303 WFNKQGMKSKSGNVVYEMLVRYGLDPLRYILMRNLVPGSDGFTTTPEDYVGRINYL 362
QY 362 ANDGLNLRNTVMINKYFDQGPAYVGVTEPDHVLAEVAKSIAADFTHMEAVDYPRA 421
DB 363 ANDGLNLRNTVMINKYFDQGPAYVGVTEPDHVLAEVAKSIAADFTHMEAVDYPRA 421
QY 422 LEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVNSHWQASIRVVAHLIPFMMETGR 481
DB 422 LEAVWTLISRTNKYIDETAPWVLDKDEALRDLQASVNSHWQASIRVVAHLIPFMMETGR 481
QY 482 AVLTOTGLEVSLENISLADFPADTVVAKGTPPIPRNLMEERIIAYIKQEGNKPAYE 541
DB 482 AVLTOTGLEVSLENISLADFPADTVVAKGTPPIPRNLMEERIIAYIKQEGNKPAYE 541
QY 542 KENPDEVELKLNKDEIKFDDPKFVIRVAEVKESKVGSDXKLLQFRLDAGDGRQIL 601
DB 542 KENPDEVELKLNKDEIKFDDPKFVIRVAEVKESKVGSDXKLLQFRLDAGDGRQIL 601
QY 602 SGIAKYPNQEVLGKQVQIVANLKPQKMKKTVSQMILSAHDGKLLIIVDPAPVNG 661
DB 602 SGIAKYPNQEVLGKQVQIVANLKPQKMKKTVSQMILSAHDGKLLIIVDPAPVNG 661
QY 662 SVIG 665
DB 662 SVIG 665

CC initiator tRNA (fMet) aminoacylation (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA (Met) = AMP +
CC diphosphate + L-methionyl-tRNA (Met).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AE006312; AAK04890.1; -
CC PIR; H86723; H86723.
CC HSP; P23395; 1A8H.
CC HANAP; MF_01228; fused; 1.
CC InterPro; IPR004495; MetG_Cterm.
CC InterPro; IPR008224; MetRS dimerising.
CC InterPro; IPR008994; Nucleic acid OB.
CC InterPro; IPR002300; tRNA-synt 1a.
CC InterPro; IPR001412; tRNA-synt 1.
CC InterPro; IPR002304; tRNA-synt met.
CC InterPro; IPR009080; tRNA-synt la_bind.
CC InterPro; IPR002347; tRNA_bind.
CC Pfam; PF00133; tRNA-synt 1; 1.
CC Pfam; PF01588; tRNA_bind; 1.
CC PIRSF; PIRSF001528; MetRS dimerising; 1.
CC PRINTS; PR01041; TRNASYNTHMET.
CC TIGRFAMS; TIGR00398; metG; 1.
CC TIGRFAMS; TIGR00399; metG_Cterm; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
CC PROSITE; PS00886; TRBD; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Protein biosynthesis; RNA-binding; tRNA-binding.
CC SITE 14 24 "HIGH" region.
CC SITE 308 312 "KMSKS" region.
CC DOMAIN 559 662 tRNA-binding.
CC BINDING 311 311 ATP (By similarity).
CC SEQUENCE 662 AA; 75777 MW; 5538AAAD66F1CB91 CRC64;
Query Match 68.8%; Score 2392.5; DB 1; Length 662;
Best Local Similarity 68.6%; Pred. No. 5.3e-135;
Matches 458; Conservative 88; Mismatches 113; Indels 9; Gaps 7;
QY 1 MSE-KNFYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGCKIQK 59
DB 1 MTENKTFYITPTIYPSGKLHIGSAYTTIACDVLARYKRLMGVDYFYLTLGLDEHGCKIQK 60
QY 60 KAEERAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDDIY 119
DB 61 KAEERAGITPQAVYDGMVAGVKELWQLLDISYDKFIRTTDDYHEKVAQVFERLLAQQDDIY 120
QY 120 LGEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQD 179
DB 121 LGEYSGWYSVDEEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQD 180
QY 180 RLVEFFKHAHPEFTTPGRNEMLRNIEPGLDLAVSRFTTGWGVPVSPNKHVVVWID 239
DB 181 WLLQYIDHDPDIQPEVRKRNEMVNNFIKPGLEDALATTSFTWGPVSPNKHVVVWID 240
QY 240 ALLNYATALGYA-QDEHGNFDFKFWGTVFHWGKQILRPHSYWPIILLMMLDVKLPDRILAH 299
DB 241 ALLNYATALGYA-QDEHGNFDFKFWGTVFHWGKQILRPHSYWPIILLMMLDVKLPDRILAH 298
QY 300 HGFVWVKQGMKSKSGNVVYEMLVRYGLDPLRYILMRNLVPGSDGFTTTPEDYVGRINYL 359
DB 299 HGFVWVKQGMKSKSGNVVYEMLVRYGLDPLRYILMRNLVPGSDGFTTTPEDYVGRINYL 358

Db 600 NHRQILSGIAEFYDPDPAALIGKKVIVANLKPRK-MRGQISQGMILSAESPEGLQIVEA 658

QY 655 DPAVNGSVI 664

Db 659 PKEMPNGAGI 668

RESULT 10

SYM_OCEIH STANDARD; PRT; 673 AA.

AC P59079;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE 05-JUL-2004 (Rel. 44, Last annotation update)

DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).

GN Name=metG; OrderedLocusNames=OB0046;

OS Oceanobacillus ihewensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OK NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H7E831 / DSM 14371 / JCM 11309;

RX MEDLINE=22220767; PubMed=12235376;

PA Takami H., Takaki Y., Uchiyama I.;

RT "Genome sequence of Oceanobacillus ihewensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments.";

RL Nucleic Acids Res. 30:3927-3935(2002).

CC -!- FUNCTION: Is required not only for elongation of protein synthesis but also for the initiation of all mRNA translation through initiator tRNA(fMet) aminoacylation (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP + diphosphate + L-methionyl-tRNA(Met).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family. MetG subfamily 2B.

CC -!- SIMILARITY: Contains 1 tRNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AP004593; BAC12002.1; -.

CC HSSP; P23395; 1A8H.

CC HAMAP; MF 01228; fused; 1.

CC InterPro; IPR004495; MetG_Cterm.

CC InterPro; IPR008224; MetRS dimerising.

CC InterPro; IPR008994; Nucleic_acid_OB.

CC InterPro; IPR002300; tRNA-synt_1a.

CC InterPro; IPR001412; tRNA-synt_1.

CC InterPro; IPR002304; tRNA-synt_met.

CC InterPro; IPR003080; tRNAasn_la_bind.

CC InterPro; IPR002547; tRNA_bind.

CC Pfam; PF00133; tRNA-synt_1; 1.

CC Pfam; PF01588; tRNA_bind; 1.

CC PIRSF; PIRSF001528; MetRS dimerising; 1.

CC PRINTS; PR01041; TRNASYNTMET.

CC TIGRFAMs; TIGR00398; metG; 1.

CC TIGRFAMs; TIGR00399; metG_C term; 1.

CC PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.

CC PROSITE; PS00886; TRSD; 1.

KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;

KW Protein biosynthesis; RNA-binding; tRNA-binding.

FT SITE 14 24 "HIGH" region.

FT SITE 310 314 "KMSKS" region.

FT DOMAIN 571 673 tRNA-binding.

FT BINDING 313 313 ATP (By similarity).

SQ SEQUENCE 673 AA; 76610 MW; D6E8F706BD9BCAD CRC64;

Query Match 64.1%; Score 2230.5; DB 1; Length 673;

Best Local Similarity 62.6%; Pred No. 2.8e-125;

Matches 420; Conservative 100; Mismatches 140; Indels 11; Gaps 6;

QY 3 EKNFYITPIYYPYSGKLHIGSAITTIACDVARYRLMGYDVYFYLTTGDEHGQKIQOARAE 62

Db 4 ENAFYITPIYYPYSGKLHIGNAITTIACDVARYRMGRGDFVYLTGSDHGQKIEQAK 63

QY 63 EAGITFOAVDGMVGVKELMQLLDSYDKFTRTDDVHEKVVAOVFERLLAQQDDIYLGE 122

Db 64 EWNISFKAYVDDMAEGMKMLWNTLEISNDKIRITTEEHQKKVVADIFERFEGQDIYDE 123

QY 123 YSGWYSVSDDEEFFETSQALAEVRDEAGNVGTGIAP-SCHEVEWVSESEYFRLSKYQRL 181

Db 124 YEGWYSVPDETETYTQLEDVERDDGNVIGKSPDSGHPVELIKEESYFFRMSKYADRL 183

QY 182 VEFFKAHPEFITPDGRLAEMLENFTERGLEDAVSTTFTWGPVPYSPKVVVYVWIDAL 241

Db 184 LKFYEDNPEFIOPESRKNEMINNF:KPGLEDLAVSRITTFSGVQVPSMKPVVYVWIDAL 243

QY 242 LNVATALGYAODEHGNFDFKFWNGTTFHVMVGKDILFFHSIYWPILMLMLDVKLPDLRIAG 301

Db 244 TNYITALGYGSEDTSLYDKFWPADV-HVMGKEIVAFHTIYWPILMALDLPLPKKVFAG 302

QY 302 WFMKDGKMSKGNVYTPPEMLVERYGLDPLRYIMRNLVPGSDGTFTFPEDYVGVINTEL 361

Db 303 WLMKDGKMSKGNVYTPPEMLVERYGLDALRYILMREVAFGSDGVFTPEDFISRVNYDL 362

QY 362 ANDLGNLNRVTSMINKYVDFGQIPAYVEGVTDFDHVLAEVAEKSIADFHHEAVDYDPA 421

Db 363 ANDLGNLNRVTAMINKYVDFGKVPFKGEVTSFDGELQTTANNAVKEKHEGQVFSDA 422

QY 422 LEAVMTLSRTNKYIDETAPWVLDKDEALRDQASVMSHQASIRVVAHLIEPFFMETS 481

Db 423 LKQWILISRANKYIDETEPWIVAXDEGRNELASVWHLAESLHAAALMLQPLTHAPK 482

QY 482 AVLTOIGLEEVSSLE--NLSLADFPADVTYVAKGTPIPRLNWEELIAYIKOMEKNKA 539

Db 483 KIAEQGLGEEYGLDWGTGFGNFENTTVKKGTPIPRLDLDEAAIYRDMANGANA 542

QY 540 VEKE-----WNPDEVELKINKD-EIKFEDFDKVEIRVAEVEKSVKVGSDKLLQFRLDAG 593

Db 543 ASSEDETGDWDFNETDLVSEKEKQIKYDFDKVLAEVDKDCSVKVGADKLLKFLRDAG 602

QY 594 DEEDQILSGIAKYVNEQELVGKQIVANLKPRKMKKYVSCGMILSAEHDGKLTLLT 653

Db 603 DNGDRQILSGIAEYYSPEQLIGKKVIVANLKPRK-MRGEISQGMILSAEYDGLQIVE 661

QY 654 VDPVAVNGSVI 664

Db 662 APSEAPNGSSI 672

RESULT 11

SYM_LACPL

ID SYM_LACPL STANDARD; PRT; 683 AA.

AC Q88Z97;

DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).

DE (MetRS).

GN Name=metG; Synonym=metS; OrderedLocusNames=lp_0454;

OS Lactobacillus plantarum.

OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;

OC Lactobacillus.

OX NCBI_TaxID=1590;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCIMS 4826 / WCPSL;

RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;

RA Kleebebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hofer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Stezen R.J.,
 RT "Complete genome sequence of *Lactobacillus plantarum* WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 CC -i- FUNCTION: Is required not only for elongation of protein synthesis
 CC but also for the initiation of all mRNA translation through
 CC initiator tRNA(fmet) aminoacylation (By similarity). AMP +
 CC -i- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 CC diphosphate + L-methionyl-tRNA(Met).
 CC -i- SUBUNIT: Homodimer (By similarity).
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -i- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC MetG subfamily 2B.
 CC -i- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL935253; CAD63085.1; -;
 CC HSP; O9V011; IMKH.
 DR HAMAP; MF_01228; -; 1.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR008994; Nucleic_acid_OB.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-synt_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind; 1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
 DR PROSITE; PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Protein biosynthesis; tRNA-binding.
 FT SITE 15 25 "HIGH" region.
 FT SITE 311 315 "KMSKS" region.
 FT DOMAIN 581 683 tRNA-binding.
 FT BINDING 314 314 ATP (By similarity).
 SQ SEQUENCE 683 AA; 76796 MW; 2E68E67D28DA8560 CRC64;
 Query Match 63.3%; Score 2201; DB 1; Length 683;
 Best Local Similarity 61.4%; Pred. No. 1.7e-123;
 Matches 416; Conservative 106; Mismatches 137; Indels 20; Gaps 7;
 QY 2 SERNFYITPTIYPYSGKLHGSAITTIACDLVARYKRLMGYDVYLTGDERGKIQKQA 61
 DB 4 TKPTVITTTIYPYSGKLHGNSYTTIACDLVARYKRMGYDVYLTGDERGLKIEKA 63
 QY 62 BEAGITPOAVDGMVGVKELWOLLISVDKFTTTDDYHKVAVQVFERLLAQDDIYLG 121
 DB 64 EKLTNDTPKSYVDGMQAKIDKWLWLLSINDKFIITDDYHRAVQVIFPRLKNGDIYUG 123
 QY 122 EYSGWTSVSDDEEFTBESQLAEVFRDAGNVGTGAGPSGHEVWVSDESYFLRLSKYQDRL 181
 DB 124 EYSGWTSVSDDEEFTETQLAEVFRDDNGKVIQKAPSGHEVELVKEQSYFFKMSKYADWL 183
 QY 182 VEFKAPHEITPDGRINLEMLNFIEPGLDLAVSTTTTGWVPSPNKHVYVWIDAL 241
 DB 184 LDYQSHPDIFIEPANRNTMINNFIPGLDLAVSTSTGWVPKSDPKHVYVWIDAL 243
 QY 242 LNYATALGYAQ-DEHGNFDFKMGVTFVHMVGKIDILRFHSIYWPILMLDLVKLPDLIAH 300
 DB 244 TNYITALGYATGSDSLFNKFWADV-QMVGKEIVRFHYIWPILHALGLPLPKKVFCH 302

QY 301 GWFVKDKGMSKSGNVVYPEMLVERGLDPLRYLMRLNLPVSGDGTFTPTDPYVGRINYE 360
 DB 303 GLLMKDKGMSKSGNVVYPTETLVERGLDALRYLVKAMPYGNGLFTPEDFVARVNYD 362
 QY 361 LANDLGNLNRVTSMINKYFDGQIPAYVEGVTFEDHVLAEVAKSADFTHEAVDYPR 420
 DB 363 LANDLGNLNRRTAMINKYEDGQIPAFKAGTVFDDADLEATAATTIKNFACNACMSLHLS 422
 QY 421 ALAEVMTLISRTNKYIDETAPWVLDDKDEALR--DQLASVMHMQASIRVVVAHLIEPFME 478
 DB 423 ALAEVWKLVSRTNKYIDETAPWQLAKSDDDANDAKLASVMAHLAASLRVIASLSPWTH 482
 QY 479 TSRAVLFOIGLEEVSLLENLSADPPADVTVVAKGTPPPRPLNMBEETAYIKEQM---- 533
 DB 483 APKEIFTQLGLDPATLAIADLQADLPAGAQQVYVAKGTPPIPRVMDMAEVEFLKGMKTS 542
 QY 534 --EGNK-----PAVEKENPDEVELKLNKDEIKFEDPKVEIRVAEVEKYSKVEGSD 583
 DB 543 KQGRKAMENAKHEAEVEQGNPAETNLTNLTPTAITIDDFDKVELKVAEVIITVKLKGAD 602
 QY 584 KLLQFRLDAGDGEDRQILSGIAKYIPNEQELVGKQVQIVANLKPRLKMKYKQGMILSA 643
 DB 603 KLLQFRLDAGDADHRQILSGIAKYWPEPEBLIGKVIIVGNLKPRLK-LRGEMSQGMILSA 661
 QY 644 BHDGKLTLLTVDPAPVNGSVI 664
 DB 662 BHDGQVQLITVPDMMVNGSLI 682
 RESULT 12
 Q724N6 PRELIMINARY; PRT; 665 AA.
 ID Q724N6 Q724N6;
 AC Q724N6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10).
 GN Name=metG; OrderedLocusNames=LMOF2365_0188;
 OS Listeria monocytogenes (serotype 4b / strain F2365).
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 CX NCBI_TaxID=265669;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15115801; DOI=10.1093/nar/gkh562;
 RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,
 RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
 RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
 RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,
 RA Forberger H.A., Tran B., Katharicu S., Wonderling L.D., Ulrich G.A.,
 RA Bayles D.O., Luchansky J.B., Fraser C.M.,
 RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
 RT borne pathogen *Listeria monocytogenes* reveal new insights into the
 RT core genome components of this species.";
 RL Nucleic Acids Res. 32:2386-2395(2004).
 DR EMBL; AE017322; AAT02975.1;
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004825; F:methionine-tRNA ligase activity; IEA.
 DR InterPro; IPR004495; MetG Cterm.
 DR InterPro; IPR008224; Nucleic_acid_OB.
 DR InterPro; IPR008994; tRNA-dimerising.
 DR InterPro; IPR002300; tRNA-synt_1a.
 DR InterPro; IPR002304; tRNA-synt_met.
 DR InterPro; IPR009080; tRNA-synt_1a_bind.
 DR InterPro; IPR002547; tRNA_bind.
 DR Pfam; PF00133; tRNA-synt_1; 1.
 DR Pfam; PF01588; tRNA_bind; 1.
 DR PIRSP; PIRSF001528; MetG dimerising; 1.
 DR PRINTS; PR01041; TRNASYNTHMET.
 DR TIGRFAMs; TIGR00398; metG; 1.
 DR TIGRFAMs; TIGR00399; metG C term; 1.
 DR PROSITE; PS50886; TRBD; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.

SQ SEQUENCE 665 AA; 75767 MW; B4FCE15778689C27 CRC64;
 Query Match 60.8%; Score 2114; DB 2; Length 665;
 Best Local Similarity 60.9%; Pred. No. 2.7e-118;
 Matches 408; Conservative 98; Mismatches 146; Indels 18; Gaps 8;
 Qy 3 EKN-FYITTIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVFLYLTGLDEHGQKIQOKA 61
 Db 5 EKNFYITTIYPSGKAHIGHAYTTVAGDAMARYKRLKGYDVFLYLTGLDEHGQKIQAKA 64
 Qy 62 EAGITPQAYVDGMAGVKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQQDIYLG 121
 Db 65 KERGISEQYVDEIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEQGDYILG 124
 Qy 122 EYSGWYVSDEBFFTESQLAEVFERDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 181
 Db 125 EYEGWYVSDEBFFTESQLAEVFERDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 184
 Qy 182 VEPFKAHPETTPDGRLEMLRNFIEPGLDLAVSRITFTWGVPSNPKHVVYVMDAL 241
 Db 185 VEYNSHPETTPDGRLEMLRNFIEPGLDLAVSRITFTWGVPSNPKHVVYVMDAL 244
 Qy 242 LNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDILRPHSIYWPILLMMLDVKLPDRLIAHG 301
 Db 245 SNYITALGYNTDNTKFKYWPADV-QIVGKEIVRPHIYWPIMLMALDLPKMWVFGH 303
 Qy 302 WFWKDGKMSKGNVYVPEMLVERYGLDPLRYVLMENLPVSGDGTFTPDYVGRINVEL 361
 Db 304 WILMKDGKMSKGNVYVPEMLIDRYGLDALRYLLREVFGSDGLFTPDYVGRINVEL 363
 Qy 362 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEFDHVLAEVAEKSIADFFTHMEAVDYPRA 421
 Db 364 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEFDHVLAEVAEKSIADFFTHMEAVDYPRA 423
 Qy 422 LEAVNTLISRTNKYIDETAPWLDKDEALRDQALASVMSHWQASIRVVAHLIEPFMMETSR 481
 Db 424 LNQLWSLISRTNKYIDETAPWLDKDEALRDQALASVMSHWQASIRVVAHLIEPFMMETSR 483
 Qy 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVVAKGTPIPPRLNMEEEIAYIKEOME 534
 Db 484 EIFLQGLQE-----ENLKKWDSIYGYGEIPAGTIVVKKGTPIFPRLEAEVEVYIQDEM 539
 Qy 535 GNKPAVEKWNPDVEVLKLNKDEIKFDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGD 594
 Db 540 GSAPAPAKE--AAEVE-ALETPOIGIEDFDKIDLRVAEVKQVDKVKKADKLLCFQLDLGE 596
 Qy 595 GEDROILSGIAKYYPNEQELVGKQVIVANLKPKRMKKYVSGMILSAEHDKGLTLTV 654
 Db 597 GKLRQVLSGIAEYFQPE-ELIGKKVIVVSNLKPVK-LRGLMSEGMILSGEKGKLSVIEA 654
 Qy 655 DPAPNGSVI 664
 Db 655 SSALPNGAKV 664
 RESULT 13
 AAT02975 PRELIMINARY; PRT; 665 AA.
 AC AAT02975;
 DT 10-MAY-2004 (Tremblrel. 27, Created)
 DT 10-MAY-2004 (Tremblrel. 27, Last sequence update)
 DT 10-MAY-2004 (Tremblrel. 27, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10).
 GN METG OR LMOF2365.0188.
 OS Listeria monocytogenes str. 4b F2365.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria;
 OC Listeria monocytogenes.
 OX NCBI_taxid=265669;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4b F2365;
 RC PubMed=15115801;
 RA Nelson K.E., Fouts D.E., Mongodin E.F., Ravel J., DeBoy R.T.,

RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,
 RA Peterson J., White O., Nelson W.C., Nierman W.C., Beanan M.J.,
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,
 RA Haft D.H., Sellenet J., Van Aken S., Khouri H., Fedorova N.,
 RA Forberger H., Tran B., Kathariou S., Wonderling L.D., Uhlir G.A.,
 RA Bayles D.O., Luchansky J.B., Fraser C.M.;
 RT "Whole genome comparisons of serotype 4b and 1/2a strains of the food-
 RT borne pathogen *Listeria monocytogenes* reveal new insights into the
 RT core genome components of this species.";
 RL Nucleic Acids Res. 32:2386-2395 (2004).
 DR EMBL; AB017322; AAT02975.1; -;
 KW Aminoacyl-tRNA synthetase; Ligase.
 SQ SEQUENCE 665 AA; 75767 MW; B4FCE15778689C27 CRC64;
 Query Match 60.8%; Score 2114; DB 2; Length 665;
 Best Local Similarity 60.9%; Pred. No. 2.7e-118;
 Matches 408; Conservative 98; Mismatches 146; Indels 18; Gaps 8;
 Qy 3 EKN-FYITTIYPSGKLHIGSAVTTIACDVLARYKRLMGYDVFLYLTGLDEHGQKIQOKA 61
 Db 5 EKNFYITTIYPSGKAHIGHAYTTVAGDAMARYKRLKGYDVFLYLTGLDEHGQKIQAKA 64
 Qy 62 EAGITPQAYVDGMAGVKELWQLLDISYDKPIRTTDDYHEKVAQVFERLLAQQDIYLG 121
 Db 65 KERGISEQYVDEIAEGFQELWKLKLEISNTDFIRTTQDRHKTSVEKIFEQLLEQGDYILG 124
 Qy 122 EYSGWYVSDEBFFTESQLAEVFERDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 181
 Db 125 EYEGWYVSDEBFFTESQLAEVFERDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 184
 Qy 182 VEPFKAHPETTPDGRLEMLRNFIEPGLDLAVSRITFTWGVPSNPKHVVYVMDAL 241
 Db 185 VEYNSHPETTPDGRLEMLRNFIEPGLDLAVSRITFTWGVPSNPKHVVYVMDAL 244
 Qy 242 LNYATALGYAQDEHGNFDFKFWNGTVFHMVGKDILRPHSIYWPILLMMLDVKLPDRLIAHG 301
 Db 245 SNYITALGYNTDNTKFKYWPADV-QIVGKEIVRPHIYWPIMLMALDLPKMWVFGH 303
 Qy 302 WFWKDGKMSKGNVYVPEMLVERYGLDPLRYVLMENLPVSGDGTFTPDYVGRINVEL 361
 Db 304 WILMKDGKMSKGNVYVPEMLIDRYGLDALRYLLREVFGSDGLFTPDYVGRINVEL 363
 Qy 362 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEFDHVLAEVAEKSIADFFTHMEAVDYPRA 421
 Db 364 ANDLGNLLNRTVSMINKYFDGQIPAYVGVTEFDHVLAEVAEKSIADFFTHMEAVDYPRA 423
 Qy 422 LEAVNTLISRTNKYIDETAPWLDKDEALRDQALASVMSHWQASIRVVAHLIEPFMMETSR 481
 Db 424 LNQLWSLISRTNKYIDETAPWLDKDEALRDQALASVMSHWQASIRVVAHLIEPFMMETSR 483
 Qy 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVVAKGTPIPPRLNMEEEIAYIKEOME 534
 Db 484 EIFLQGLQE-----ENLKKWDSIYGYGEIPAGTIVVKKGTPIFPRLEAEVEVYIQDEM 539
 Qy 535 GNKPAVEKWNPDVEVLKLNKDEIKFDPKVEIRVAEVEKSVKVGSDKLLQFLRDAGD 594
 Db 540 GSAPAPAKE--AAEVE-ALETPOIGIEDFDKIDLRVAEVKQVDKVKKADKLLCFQLDLGE 596
 Qy 595 GEDROILSGIAKYYPNEQELVGKQVIVANLKPKRMKKYVSGMILSAEHDKGLTLTV 654
 Db 597 GKLRQVLSGIAEYFQPE-ELIGKKVIVVSNLKPVK-LRGLMSEGMILSGEKGKLSVIEA 654
 Qy 655 DPAPNGSVI 664
 Db 655 SSALPNGAKV 664
 RESULT 14
 SYM_LISMO
 ID SYM_LISMO STANDARD; PRT; 664 AA.
 AC 08YAF2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

	Matches	407; Conservative	101; Mismatches	144; Indels	18; Gaps
QY	3	EKN-FYITPIYPSGKLHIGSAYTTIACDVLARYKRLMGVYFYVYLTGLDGHGKIQQA	61		
Db	4	EKNFYITPIYPSGKAHICHAYTTVAGDAMARYKRLGVDYFYLTGLDGHGKIQQA	63		
QY	62	EAAGITQAYVDGAVGVKELWQLLDISYDKFIRTTDDVHEKVVAAVQRELLAQQDIYLG	121		
Db	64	KERGISQGYVDEIABQFQELWKKLEISNTDFIRTTQDRHKTSVEKIFQGLEQGDYIYG	123		
QY	122	EYSGWYSVDSDEFFTESQLAEVPRDEAGNVTGGIAPSGCHEVWVSBEYSYFLRLSKYQDL	181		
Db	124	EYEGWYSVDSDEYFTETQLEEVYKDKNGKVIQGGKAPSGNEVELVKESYFFRMSKYADRL	183		
QY	182	VEFFKAMPERTTPGRNLNEMLRPIERGLEDLAVSRRTFTWGVYPSPNPKHVYVYWDAL	241		
Db	184	VEYNSHPEFTLPSRKNEMINPIKPLGSLDLAVSRRTFDWGIKVPNGPHVYVYWDAL	243		
QY	242	LNATACGYAQDEHGNFDFKFWNGTVFHMVGKDLIRFHSIYWPILLMMLDVKLPDLRIAH	301		
Db	244	SNYITAGYNTDNTKQKYWPADV-QIVGKEIVRFTIYWPIMLXALDLPLPKWVEGHG	302		
QY	302	WFWKDGKMSKSGNVVYPEWLVERYGDLPIRYILMRNLVPGSDGTFTEPDYVGRINYL	361		
Db	303	WILKDGKMSKSGNVVDPYMLIDRYGLDALRYLLREVPFGSDGLTFTEPDFVDRYNFDL	362		
QY	362	ANDLGNLLNRVTSMINKYFDGQIPAYVEGTEFDHVLAEVAEKSIADFTHTMEADVYRA	421		
Db	363	ANDLGNLLNRVAMINKYFDGEIPAYQGVNTEFDQLIVDFKNVVKESGMDHMQFSVA	422		
QY	422	LEAVWTLISRTNKYIDTAPWLVKDQALBDQLASVNSHWQASIRVVVAHLIIPFMMETR	481		
Db	423	LNQLWSLISRTNKYIDETAPWALAKDBDKETELASVNTHTLAENLRITIAVLLOPFLTRTPG	482		
QY	482	AVLTQIGLEEVSSLENI-----SLADFFADTVTVAKGTPPIPRPLNMBEETAYIKEOME	534		
Db	483	EIFQLGLQE-----ENLKKWDSIYGYEIPAGTIVVKKTPPIPRLDAEVEVYTIQDENK	538		
QY	535	GKPKAVEKEMNPDEVELKANKDBIKFEDFDKVBIRVAEVKESKVBEGSKLQFLRLDAGD	594		
Db	539	GSAPAPAE--VAEVE-ALETPOIGIEDFDKIDLRVAEVKQVDKVKAKDKLCLCFQLDLGE	595		
QY	595	GEDQQLISGIAKYYPNQELVGKKVQIVANLKRPMKKKVVSGMTLSAEHDKLTLVV	654		
Db	596	GKLQVLSGIAEYFQPS-ELIGKKVIVVSNLKPVK-LRGLMSGMTLSGKDGKLSVIEA	656		
QY	655	DPAPVNGSVI	664		
Db	654	SSALPNGAKV	663		
RESULT 15					
SYN_LISIN	ID	SYM	LISIN	STANDARD;	PRT; 664 AA.
AC	Q92F90;				
DT	28-FEB-2003	(Rel. 41, Created)			
DT	28-FEB-2003	(Rel. 41, Last sequence update)			
DE	05-JUL-2004	(Rel. 44, Last annotation update)			
DE	Methionyl-tRNA synthetase (EC 6.1.1.10)	(Methionine--tRNA ligase) (MetRS);			
DE	Name=metG; Synonym=metS; OrderedLocusNames=lin0216;				
OS	Listeria innocua				
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.				
OX	NCBI_taxid=1642;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CLIP 11262 / Serovar 6a;				
RC	MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;				
RA	Glaser P., Franeul L., Buchrieser C., Rusniok C., Amend A.,				
RA	Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,				
RA	Charbit A., Cherouani F., Couve E., de Daruvar A., Dehoux P.,				
RA	Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dueserget O.,				
RA	Entian K.-D., Fishi H., Garcia-del Portillo F., Garrido P.,				

```

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Matounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstrek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RL "Comparative genomics of Listeria species.";
SC Science 294:849-852(2001).
CC -i- FUNCTION: Is required not only for elongation of protein synthesis
CC but also for the initiation of all mRNA translation through
CC initiator tRNA(fMet) aminoacylation (By similarity). = AMP +
CC -i- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC diphosphate + L-methionyl-tRNA(Met).
CC -i- SUBUNIT: Homodimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -i- SIMILARITY: Belongs to Class-I aminoacyl-tRNA synthetase family.
CC MetG subfamily 2B.
CC -i- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL596163; CAC95449.1; .
CC PIR; A11459; A11459.
CC HSP; P21888; 1117.
CC Listlist; LIN00216; -.
CC HAMAP; MF_01228; fused; 1.
CC InterPro; IPR004495; MetG_Cterm.
CC InterPro; IPR008224; MetRS dimerising.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002304; tRNA-synt_met.
CC InterPro; IPR003080; tRNA-synt_1a_bind.
CC InterPro; IPR002547; tRNA_bind.
CC Pfam; PF00133; tRNA-synt_1; 1.
CC Pfam; PF01588; tRNA bind; 1.
CC PIRSF; PIRSF001528; MetRS dimerising; 1.
CC PRINTS; PR01041; TRNASYNTHMET.
CC TIGRFAMs; TIGR00398; metG; 1.
CC TIGRFAMs; TIGR00399; metG_Cterm; 1.
CC PROSITE; PS00178; AA TRNA_LIGASE_I; FALSE_NEG.
CC PROSITE; PS00886; TRBD; 1.
CC Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
CC Aminoacyl-tRNA synthetase; RNA-binding; tRNA-binding.
CC Protein biosynthesis; RNA-binding.
CC SITE 15 25 "HIGH" region.
CC SITE 310 314 "KXSKS" region.
CC DOMAIN 563 664 tRNA-binding.
CC BINDING 313 313 ATP (By similarity).
CC SEQUENCE 664 AA; 75756 MW; 64A526B817BB195 CRC64;
Query Match 60.6%; Score 2107; DB 1; Length 664;
Best Local Similarity 60.6%; Pred. No. 7e-118;
Matches 406; Conservative 101; Mismatches 145; Indels 18; Gaps 8;
QY 3 EKN-FYITPTIYPSGKLHGSAVTTIACDLARYKLMGYDVPYLTGLDEHGQIOCKA 61
DB 4 EKNTFYITPTIYPSGKAHGHAYTTVAGDAMARYKRLKGYDFYLTGTDHGQIOAKA 63
QY 62 BEAGITPAVYDGMAGVKEWKLQLLDISYDFIRTTDDYHEKVAQVFERLLAQDDIYLG 121
DB 64 KERGISQEYVDEIATGFCLEWKKLEISNTFIITQDRKTSVEKIFEQLEGGDIYLG 123
QY 122 EYSGWYSVDSDEFFTESQLAEVFRDEAGNVTGGIAPSGHEVWVSESYFLRLSKYQDRL 181
DB 124 EYEGWYSVDSDEYFTETQLEEVYKDENGKVIGGKAPSGNEVELVKESYFFRNSKYADRL 183
QY 182 VEFKAPHEFTPTDGRINMLRNFIKGLDLAVSRFTTFDWGIKVPGNPKHVYVWIDAL 241

```

```

DB 184 VEYNSHPFEFILPSRKNEMINNFIKPGLEDLAVSRFTTFDWGIKVPGNPKHVYVWIDAL 243
QY 242 LNYATALGYAODEHGNFDFKFWNGTVFHMVGXDIIRFHSIYWPILIMLMDVKLPORLIAGH 301
DB 244 SNYITALGYNTDNDTKFKQYWPADV-QIVGKEIIVRFHTIYWPIMLMALDLPKXWFGHG 302
QY 302 WFMKDGKXWSKSGNVVPEMLVRYGLDPLRYLYMNLNLPVSGDGTTFPDIYGRINVEL 361
DB 303 WILMKDGKXWSKSGNVVDPYMLIDRYGLDALRYLLAEVFPFGDGLTFPFDVDRVNDL 362
QY 362 ANDLGNLNRVTSMINKYFDGCIIPAYVEGVTEFDHVLAEVAKESIAOPHTHMEAVDYPRA 421
DB 363 ANDLGNLNRVTSMINKYFDGCIIPAYVEGVTEFDHVLAEVAKESIAOPHTHMEAVDYPRA 422
QY 422 LEAVWTLISRTNKYIDETAPVWLKDEALRDQLASVMSHMQASIRVVAHLIEPFWMTSR 481
DB 423 LNQLWSLISRTNKYIDETAPWALAKEEDKRTELASVMTHLAENLRIIAVLLQPLTRTPG 482
QY 482 AVLTOIGLEEVSSLENL-----SLADFPADVTVWAKGTPIFPPLNNEEERAIYKEQME 534
DB 483 EIFLQLGLQE-----ENLKKWSIYGYGPIPGTTVVKGTTFIPALDRAEVEVYIQDEM 538
QY 535 GNKPAVEKEWNPDEVELKLNKDEIKFEDFKVEIRVAEVEKSVKVGSDKLLQFLDAGD 594
DB 539 GSAPAPAE--TAEVE-ALETPQIGIEDFKIDLRVAEVKQVDKVKADKLLCFOLDLGE 595
QY 595 GEDROILSGIAKYVNEQELVKKVQIVANLKPRKMKKYVSQGMILSAEHDGKLTILTV 654
DB 596 GKLRQVLSGIAEFYQPE-ELIGKKYIVVSNLKPVK-LRGLMSEGMILSGEXDKGKLSVIEA 653
QY 655 DPAVPNGSVI 664
DB 654 NSALPNGAKV 663

```

Search completed: December 17, 2004, 21:59:30
Job time : 204 secs

This Page Blank (uspto)